AD_____

GRANT NUMBER DAMD17-96-1-6254

TITLE: Computer-Aided Diagnosis and Feature-Guided Data Reduction Systems In Mammography

PRINCIPAL INVESTIGATOR: Heang-Ping Chan, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, MI 48103-1274

REPORT DATE: October 1998

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1998	Annual (23 Sep 97 - 22 Sep 98)

4. TITLE AND SUBTITLE
Computer-Aided Diagnosis and Feature-Guided Data Reduction Systems In Mammography

5. FUNDING NUMBERS
DAMD17-96-1-6254

6. AUTHOR(S)
Heang-Ping Chan, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, MI 48103-1274

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

The goals of this project are: (1) Implement and evaluate a computer-aided diagnosis (CAD) system to assist radiologists in mammographic interpretation. (2) Develop and evaluate a feature-guided data compression technique to facilitate implementation of digital mammography.

For the first sub-project, we plan to implement our CAD algorithms for detection and classification of microcalcifications in a high speed workstation, develop user interface for efficient operation of the CAD programs, and perform a pilot clinical trial. For the second sub-project, we plan to evaluate and select the most efficient lossless and lossy data compression techniques that can provide maximum compression ratio without noticeable loss of information for mammography.

In the second year of the funding period, we have performed the following studies: (1) completion of the graphical user interface (GUI) development based on a PC which is networked to the CAD system, (2) improvement and implementation of the mass detection program in the CAD system, (3) continue the improvement and implementation of the microcalcification detection program, (4) quantitative evaluation of the data compression methods for mammograms, and (5) design methods and programs for data collection and data analysis for the pilot clinical trial. These studies are the necessary steps to accomplish the goals of this project.

14. SUBJECT TERMS Breast Cancer
Computer-aided diagnosis, data compression, breast cancer detection

15. NUMBER OF PAGES
39

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

__✓__ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

__✓__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Chan Heang Ping 10/26/98

PI - Signature Date

(4) Table of Contents

(5) Introduction

In the United States, breast cancer is the leading cause of death in women between 40 to 55 years of age(1990). It is estimated that one out of eight women will develop breast cancer in their lifetime (Boring, et al. 1994, Harris, et al. 1992). There is considerable evidence that early diagnosis and treatment significantly improves the chance of survival for patients with breast cancer (Byrne, et al. 1994, Curpen, et al. 1995, Feig and Hendrick 1993, Moskowitz 1987, Seidman, et al. 1987, Smart, et al. 1995). The American Cancer Society — National Cancer Institute Breast Cancer Detection Demonstration Project (BCDDP) has shown that mammography contributes significantly in the detection of localized breast cancer in asymptomatic women (Seidman, et al. 1987).

Although mammography has a high sensitivity for detection of breast cancers when compared to other diagnostic modalities, studies indicate that radiologists do not detect all carcinomas that are visible on retrospective analyses of the images (Baines, et al. 1986, Bassett, et al. 1987, Bird, et al. 1992, Harvey, et al. 1993, Haug, et al. 1987, Hillman, et al. 1987, Kalisher 1979, Martin, et al. 1979, Moskowitz 1987, Wallis, et al. 1991). While double reading can reduce the miss rate in radiographic reading (Metz and Shen 1992, Thurfjell, et al. 1994), it also increases the cost of screening. In our ROC study (Chan, et al. 1990), we found that a CAD scheme, which alerts the radiologist to suspicious clusters of microcalcifications, can significantly improve radiologists' accuracy in detecting the microcalcifications under experimental conditions that simulate the rapid interpretation of screening mammograms. More recently, Kegelmeyer et al. (Kegelmeyer, et al. 1994) also showed that CAD can improve radiologists' detection of spiculated masses. These studies indicate that CAD is a viable alternative to double reading by radiologists.

Early breast cancers are often characterized by subtle clustered microcalcifications and masses (Tabar and Dean 1985). It has been reported that between 30 and 50% of breast carcinomas detected radiographically demonstrate microcalcifications on mammograms, and 40 to 50% of breast carcinomas present as masses. The high correlation between the presence of microcalcifications and masses and the presence of breast cancers indicates that an increase in the accuracy of detection and analysis of the characteristic features of these lesions may lead to further improvement in the efficacy of mammography as a screening procedure for the detection of early breast cancer.

In the past few years, we have been developing CAD algorithms in detection and classification of microcalcifications and masses using advanced image processing and computer vision techniques. Our CAD algorithms have provided very promising results in laboratory tests. At this stage, it is necessary to test the algorithms in a clinical trial with a large number of mammograms obtained from the general patient population before specific methods can be developed to further improve their performance. Therefore, our goals in this proposal are to implement our CAD algorithms in a fast workstation, develop user interfaces for efficient operation of the CAD programs, and conduct a pilot clinical trial of the CAD schemes at three mammographic screening sites. Based on the results of the pilot clinical trial, we can evaluate the sensitivity and specificity of the CAD algorithms, analyze the effects of the CAD schemes on mammographic screening, identify any potential problems in a clinical environment, and develop methods to further improve the CAD schemes in the future. We believe that this is a crucial step to develop a clinically practical CAD workstation.

It has been recognized that digital mammography is one of the key research areas for improvement in the diagnosis of breast cancer (Shtern, et al. 1995). Two of the major issues in digital mammography are the technological requirements in developing high resolution digital detectors and the transmission and archiving the large amount of data. A number of solid-state large-area digital detectors

are being developed for mammographic application. It has been generally recognized that a pixel size of no greater than 0.05 mm × 0.05 mm will be required for imaging the subtle features of microcalcifications. At this resolution, a single 8" × 10" mammogram will result in 40 MB of digital data.

Data compression can reduce the amount of data for transmission and storage. However, there is often a tradeoff between compression ratio and image fidelity. Data compression in mammography is especially difficult because of the very subtle image details such as microcalcifications and mass margins that need to be preserved. We have investigated the effects of data compression on computerized detection of microcalcifications previously. In the current proposal, we plan to develop a CAD guided data compression technique to maximize the compression efficiency with a minimum loss of information. Our approach is to preserve the original image information by lossless compression in potentially important regions on the mammograms indicated by the CAD programs. For breast areas outside these regions, we will apply the most efficient lossy compression technique that does not cause noticeable degradation of image details. We will conduct both receiver operating characteristic studies and subjective image quality ranking studies to compare observer performance on the uncompressed images, on images compressed with the selected lossy technique, and on images compressed with the standard JPEG technique.

The importance of this research is based on the fact that x-ray mammography is, at present, the most reliable diagnostic procedure for detection of early breast cancer. Our proposed research aims at the development of a CAD workstation which may assist radiologists in screening and characterizing abnormalities on mammograms and the development of an efficient CAD-guided data compression technique for digital mammography. The CAD workstation, once developed, can be implemented and operated cost-effectively in various breast imaging facilities as a second opinion, and thus will potentially increase the diagnostic accuracy of mammography for breast cancer detection. The data compression technique will facilitate the implementation of telemammography and digital mammography for breast cancer screening. These new technologies therefore are expected to have a significant impact on patient care, especially in rural and remote areas.

With the support of this grant from the USAMRMC Breast Cancer Research Program, we have been developing a CAD workstation with a proper graphical user interface for a pilot clinical trial. We also continue to improve our mass and microcalcification detection programs before implementation in the CAD workstation. We are preparing cases for a subjective image quality comparison experiment to evaluate the feature guide data compression technique. Statistical methods are being developing for analysis of the pilot clinical data. We will discuss the details of these progresses in the following section.

(6) Body

During the funding period of 9/22/97 to 9/21/98, the three collaborating institutions in this Demonstration Project: University of Michigan, Georgetown University, and University of Iowa, have conducted the following tasks. The report from each of the institution is presented separately. A summary that links the tasks together and discusses the overall progress of the project is presented after the individual reports.

University of Michigan

(a) Development of a graphical user interface (GUI) for the CAD workstation

Overview of the CAD system

The CAD System (CADS) is designed to automatically process mammograms and screen the digitized images for suspicious breast lesions. At present, we are implementing two detection algorithms that can search for masses and microcalcifications on mammograms. Our long-term goal is to have a CAD system that will assist radiologists in detecting breast cancers. The current system is designed for a pilot clinical trial to evaluate the effects of the CAD system on radiologists' mammographic interpretation in a screening setting.

The structure of CADS is shown in Fig.1 (p. 12). It consists of four components: digitization of mammograms, detection of masses, detection of microcalcifications, and visualization of detection results combined with the collection of radiologist feedback. The digitization and visualization are interface modules whereas the mass detection and the microcalcification detection are processing modules. The implementation of the CADS is shown in Fig.2. A mammographic case is checked into the CADS by a barcode reader. The mammograms are then digitized with a Lumisys laser film scanner. The scanner is controlled by a personal computer (PC). After the mammograms are digitized, the images are transferred and stored in a clinical database on an optical jukebox. The mass and microcalcification detection programs are running on two separate UNIX workstations. A control program running on both UNIX workstations continuously searches for new images being stored in the jukebox. When a new image appears, this control program will initiate the execution of the mass and microcalcification detection programs on that image and send the detection results back to the jukebox for storage. The visualization program runs on the PC.

A radiologist reading clinical mammograms will first log into the CAD visualization program with a password. The radiologist will then have access to the digitized images and the CAD detection information on the jukebox. When a clinical case that has undergone CAD processing comes up on the alternator, the radiologist scans the patient barcode into the PC, the visualization program will automatically transfer and display a low-resolution version of the appropriate patient images along with the CAD information to the PC. The radiologist can then use this CAD information during clinical evaluation of the patient films. To collect data for our pilot clinical trial, the program also allows the radiologist to mark the location of any visible masses or microcalcification clusters on the images, along with action rating of the case based on the Breast Imaging Reporting and Data System (BI-RADS) scale. The following sections include more detailed description of the individual components in the CADS. Note that the names and registration numbers used in the figures do not belong to the actual patients. They were created to illustrate how the system operates.

It may be noted that we have made a major change in the design of the CAD workstation. Our original plan was to develop the GUI on the workstation, which will therefore be used both for image processing and display of the computer detection results to the radiologists. This year we have decided that a more practical approach is to display the computer output on the PC while the image processing can be centralized at any UNIX workstations available in our laboratory. All the input and output information will be stored in the optical jukebox on the network. The PC is the most convenient platform for the visualization of the computer output because all reading rooms in our hospitals and clinics are equipped with a PC. We have therefore developed a new GUI based on the PC platform, as described below.

Digitization of Mammograms

The digitization module consists of a Lumiscan 85 scanner controlled by a PC with a Pentium II 300 MHz processor (Fig.2). The films are digitized at a pixel size of 50 micron x 50 micron. A graphical user interface (GUI) was developed to streamline the digitization process. This GUI allows the operator to enter the patient information into the CAD database, and digitize and display the acquired images. Fig. 3 shows examples of the digitization GUI windows. Initially a database must be selected (either Clinical or Lab), which determines how the digitized images will be processed. Then the patient is checked in by using a barcode reader to acquire the patient registration number. In case that the patient already exists in the database the personal information and previously digitized films appear; otherwise, the operator enters the patient information. In addition, the scanning parameters may also be adjusted, if necessary, by using the **Set SCAN Parameters** menu. Parameters include the pixels per inch, bits per pixel, file format, and FTP transfer options. The FTP Options determine to which directory the image will be transferred in the Jukebox (Hewlett Packard 400EX Magneto-Optical Jukebox, 390 GB storage capacity). The destination directory changes dependent on the selected database. Images can be stored as DICOM, Lumisys or TIFF formats. In the next step, the film is scanned using the **SCAN to a File** menu. The patient personal information automatically appears when the patient's barcode has been scanned. The operator then enters the breast side and view information, inserts the film into the digitizer and acquires the image. A uniquely coded image file name is automatically generated by the program using the patient and film information. The file names of previous images (in case they exist) also appear in this window. If the scanning is successful, the digitized image is saved on the local PC disk and also displayed on the screen (Fig.4). The operator can inspect the digitized image for orientation and artifacts. If the orientation is incorrect, they can flip the image to a desire orientation by pressing a button. If the operator is satisfied with the digitized image, he/she can press the **Send** button in the FTP window and the image file will be transferred to the Jukebox for storage.

The UNIX workstations use a control program to query the Jukebox directory for new images. If a new image is found, the control program will automatically initiate the execution of the microcalcification and mass detection programs on the new image.

Detection of Masses

The mass detection module is implemented on a UNIX workstation (Fig. 2). The module consists of the five stages shown in Fig. 5. Initially the breast boundary is detected from the digitized mammogram. Segmentation of suspicious structures based on density-weighted contrast enhancement is applied to the defined breast region. Each detected object then undergoes region growing to improve the initial object borders. Eleven morphological and 32 texture based features are calculated for each of the detected structures. These features are subsequently used to differentiate between breast masses and

normal breast structures. On average, 3 to 4 regions per film will be identified by the mass detection program. Finally, the coordinates and outlines of all detected objects are saved in files and transferred back to the Jukebox for storage. These files of detection results will be accessed by the visualization program during mammographic interpretation with CAD.

Detection of Microcalcifications

The detection of the microcalcification is also carried out on a UNIX workstation (Fig. 2). Figure 6 illustrates the general scheme used to detect microcalcification clusters. The breast region of the digitized mammogram is processed with spatial filters to obtain the signal-enhanced and signal-suppressed images. A difference image is then obtained by subtracting the signal-suppressed image from the signal-enhanced image. Since the low-frequency structured background is similar in the two images, the difference image technique removes the slowly varying background from the difference image. An adaptive gray-level thresholding technique is then applied to the difference image in order to isolate a microcalcification from the remaining noise background. The resulting threshold image contains groups of pixels with values above the threshold superimposed on a uniform background. Potential microcalcifications are identified in the threshold image using an area-thresholding criterion which eliminates random noise points with areas smaller than a preselected number of pixels. Additionally a convolution neural network (CNN) trained to recognize true microcalcification patterns is used to reduce false positives (FPs). Finally a clustering criterion is used to identify microcalcification clusters containing more than a preselected number of detected microcalcifications within a predefined diameter. When detection is completed, the locations of the microcalcifications and clusters are saved in files and are transferred to the Jukebox.

Visualization of Detection Results

A GUI was developed to visualize the mass and microcalcification detection results on a PC located in the clinical reading room. Figs 7 and 8 show some of visualization GUI screens. The display screen is divided in two parts: the image display area and the information display area. The image display area shows the different digitized mammographic views along with the mass and microcalcification detection results. The information display area controls the GUI and displays patient and image information. This area is organized using tab headers to allow patient information (Fig. 7a), CAD display information (Fig.7b), local and global image windowing (Fig.8a) and display configuration information (Fig.8b) to be quickly and easy accessed. In the examples shown in Fig 7 and Fig 8, the image display area is configured to display the most recent craniocaudal (CC) and mediolateral (MLO) views of the left and right breast.

During mammographic interpretation, the radiologist will use a barcode reader to enter the patient registration number to the program. All images and CAD results associated with the patient will then be automatically downloaded from the Jukebox to the local PC disk. The images corresponding to the CC and MLO views as well as the patient information will be displayed (Fig. 7a). By clicking the **CAD ON** button, the detection results will be displayed as either mass outlines or arrows pointing to possible mass locations (Fig.7b). In the case of microcalcification detection, either the individual microcalcifications or the estimated cluster locations or both can be displayed (Fig. 8a,b). For better visualization, each of the individual image windows can also be magnified (Fig. 8b) by clicking the **zoom** button. However, this function is secondary because the radiologist will always use the original high-resolution screen-film mammograms for interpretation.

Radiologist Feedback

For the pilot clinical trial, in order to record the radiologist's detection result before and after using the computer aid, a radiologist feedback dialog has been included in the visualization GUI module. The collection sequence of the radiologist's detection results is shown in Fig. 9. The radiologist will be asked to identify any mass or microcalcification cluster location in all the views (Fig. 10a) and also give ratings for the particular case (Fig. 10b) before and after using the computer aid detection. This information is then saved in the database.

CADS Database

The CADS database contains patient information, film digitization information and radiologist feedback information. The digitization, visualization and feedback user interfaces have access to the database for reading and writing. The database is implemented in Microsoft Access and contains a total of five tables. The PatientInfo and ClinicalImageInfo Tables (Fig. 11a) are updated by the Digitization GUI program (Fig.3). The **PatientInfo** Table only contains the patient's personal information. The **ClinicalImageInfo** Table contains the scanning image information for the digitized mammograms. The MassImageObjects, CalcImageObjects and ImageQuestions Tables (Fig. 11b) are created by the radiologist feedback stage (Fig 9 and Fig 10) in the visualization GUI program. The **MassImageObjects** and **CalcImageObjects** Tables contain the coordinates of the radiologists' marked masses (Fig.10a) and microcalcifications, respectively, for each film before and after using the computer aid. The **ImageQuestions** Table contains the radiologists' rating responses shown in Fig. 10b also before and after using the computer aid. The CADS database is kept confidential by user passwords. Only the researchers and radiologists involved in the project have access to the database.

(b) Automated microcalcification detection program

As mentioned in our annual report last year, we have been converting our CAD software from the proprietary VMS operating system of the Digital Equipment Corporation (DEC) to the more portable UNIX operating system. Many of the programs have to be modified and tested because there are some differences in the FORTRAN and C compilers under the two operating systems. Most of the modifications are relatively minor. However, some of the FORTRAN programs need major changes because the original versions incorporated subroutine utilities that are specific to DEC VMS operating systems. The conversion of the mass detection programs was completed last year. The conversion of the microcalcification detection programs has been completed recently. We are currently modifying the programs to automate the entire process, upon initiation of the program execution by the query control program described above. We are also testing the performance of the microcalcification detection on randomly selected unknown cases.

(c) Automated mass detection program

The mass segmentation method has been altered during the past year to improve the borders of the detected objects and to reduce the complexity of the overall algorithm. The block diagram for the proposed detection scheme is shown in Fig. 12. Global density-weighted contrast enhancement (DWCE) segmentation is still used to identify an initial set of breast structures on the digitized mammograms. These objects are then used as starting locations for a clustering-based region-growing algorithm. The false-positive (FP) reduction techniques, which are used to differentiate between masses and normal breast structures, have been simplified in the current implementation. FP reduction is now applied to only the final set of grown objects in two stages. An initial reduction stage based on

morphological features extracted from the detected objects is followed by a texture feature based reduction stage. Previously, FP reduction was applied after the DWCE segmentation stage as well as after all region-growing stages.

The initial DWCE segmentation step employs an adaptive filter to enhance the local contrast and accentuate mammographic structures in the image. The filter is applied to the entire image on a pixel-by-pixel basis. After contrast enhancement, Laplacian-Gaussian edge detection is applied and all enclosed objects are filled to produce a set of detected structures for the image. The DWCE stage has been found to be effective in detecting most breast structures including over 90% of the breast masses. However, the DWCE borders usually fall well inside the true borders of an object and a significant number of neighboring structures are merged into single objects.

In order to improve the object margins and reduce the effects of merging, clustering-based region growing is applied to the DWCE objects. This is accomplished in two steps. First, an initial set of seed objects are determined by identifying all local maxima in the original gray-scale image which occur inside a DWCE object. In simple terms, a pixel is a local maximum if and only if its value is at least as large as all nearest neighbor pixel values. These initial maxima are expanded as follows. Gaussian smoothing $(\sigma = 2.0)$ is applied to the gray-scale image, and maximum and minimum pixel value thresholds are defined for a local maximum. All pixels within a radius of 20 pixels from a local maximum and with a pixel value inside the appropriate range are considered to be part of the object. This is repeated for all maxima within the image. The second step is then to apply K-means clustering to background-corrected regions of interest (ROIs) defined by each object. The feature images used to control the clustering consist of a median-filtered and two edge-enhanced versions of the ROI along with the original region. Clustering usually produces better border estimates than the original DWCE segmentation stage with a reduction in merging between adjacent structures.

The DWCE segmentation and growing do not differentiate masses from normal tissues, therefore, a large number of breast structures are usually detected in each mammogram. Since the shape and texture of mass objects, in general, should be different from those of normal breast structures, a set of features is extracted from each detected object and used to differentiate between the detected structures. The features are used in a sequential classification scheme to reduce the number of FP detections in a mammogram. A classifier employing 11 morphological features is initially used to eliminate objects that had shapes significantly different from breast masses. Texture features are then computed for all remaining structures and used with a linear classifier as a final arbiter between potential masses and normal structures.

We have compared the performance of this clustering-based segmentation method with our previous gradient-based method. For a data set of 253 mammograms each containing a biopsy-proven mass, both method had an initial sensitivity of over 97% following DWCE segmentation. Morphological FP reduction after clustering in comparison with morphological FP reduction applied after gradient-based region growing reduced the number of detected objects from 37 to 29 per image. The final FROC performance after texture classification was also improved with the clustering technique. At a sensitivity of 80% clustering reduced the number of FPs per image to 1.3 as compared to 1.9 FPs per image with the gradient-based growing approach. The overall free-response receiver operating characteristic (FROC) curves for both techniques are shown in shown in Fig 13. The results summarized are the test performance achieved with a group jackknife method using a 9-to-1 training-to-test ratio.

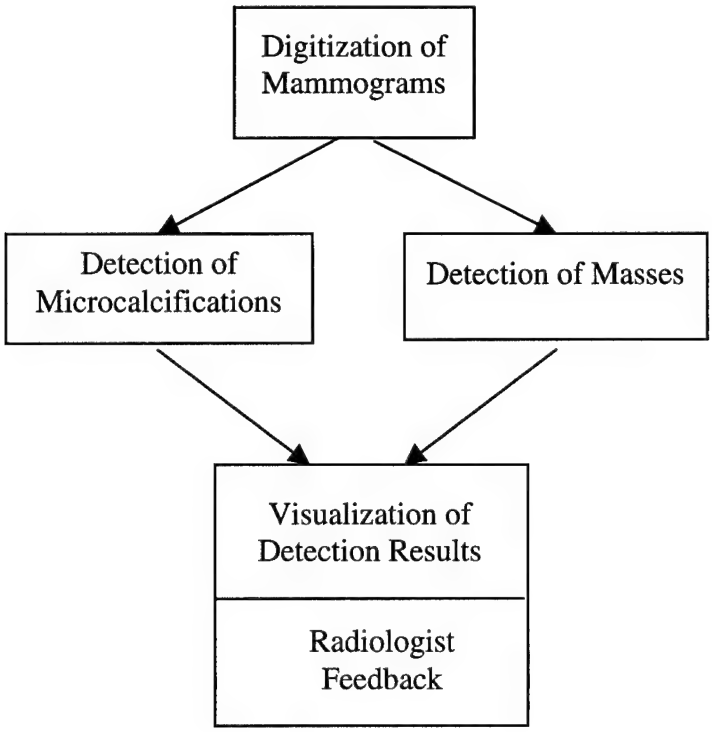


Figure 1. Structure of CAD clinical system.

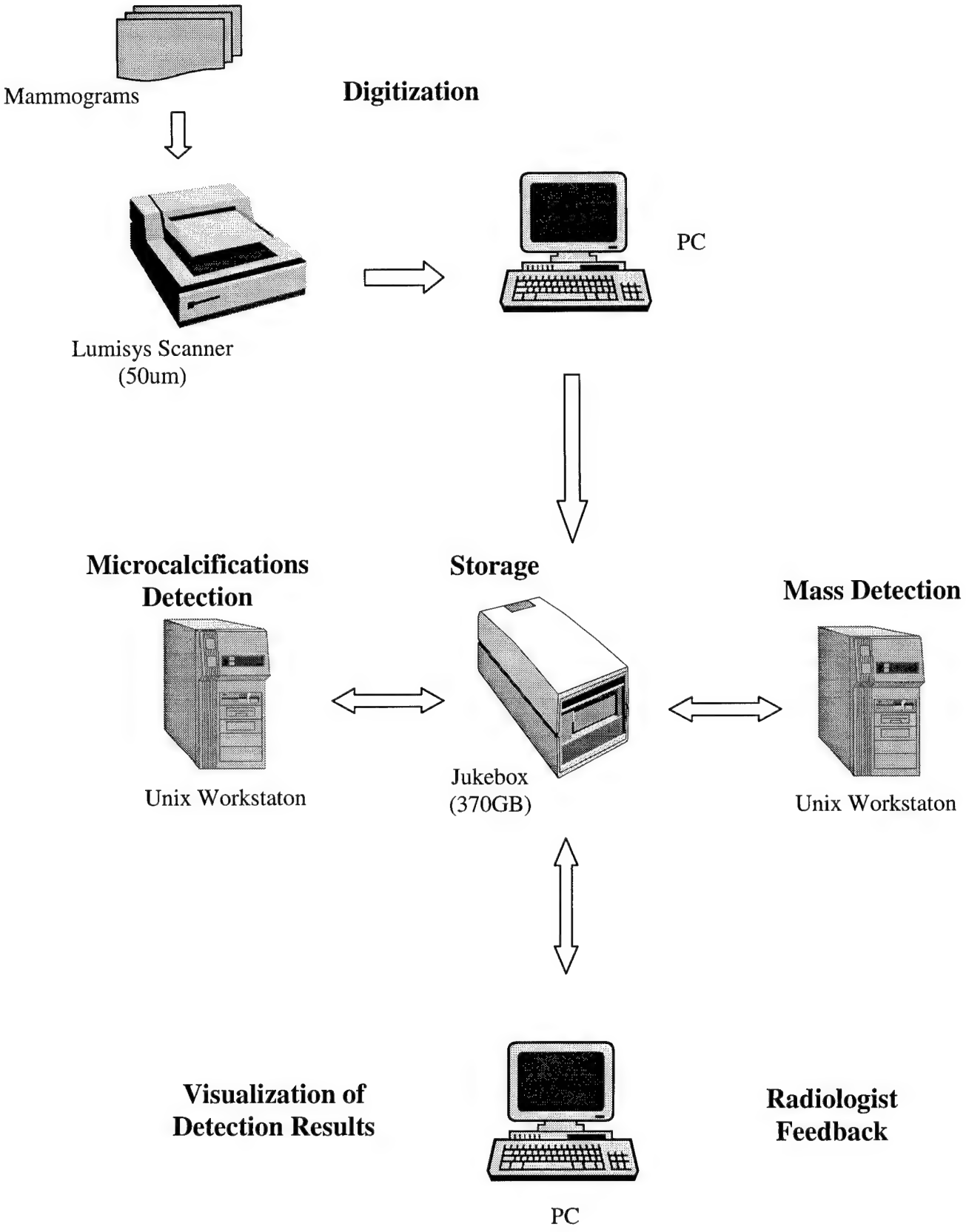


Figure 2. Implementation of CAD clinical system.

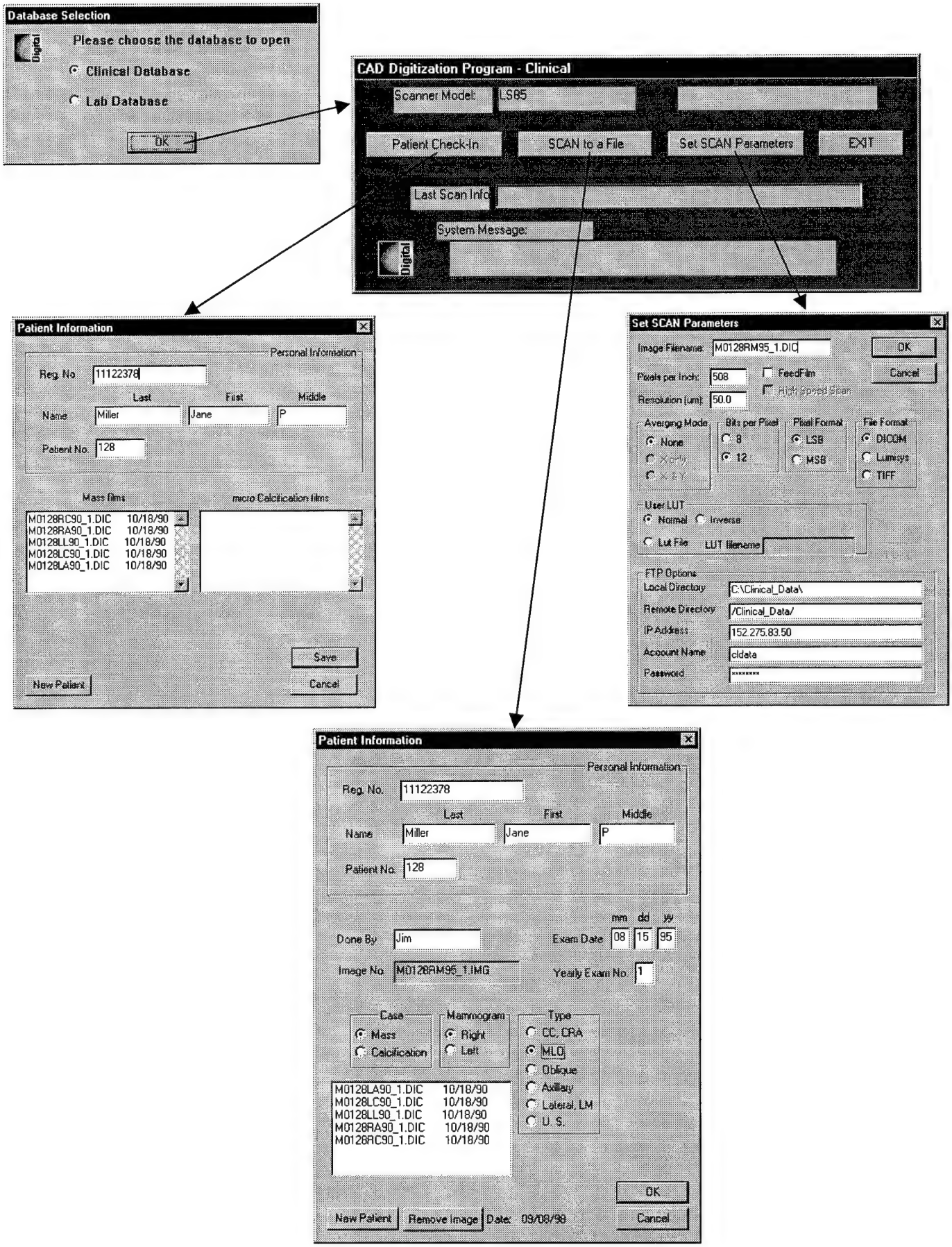


Figure 3. Digitization user interface.

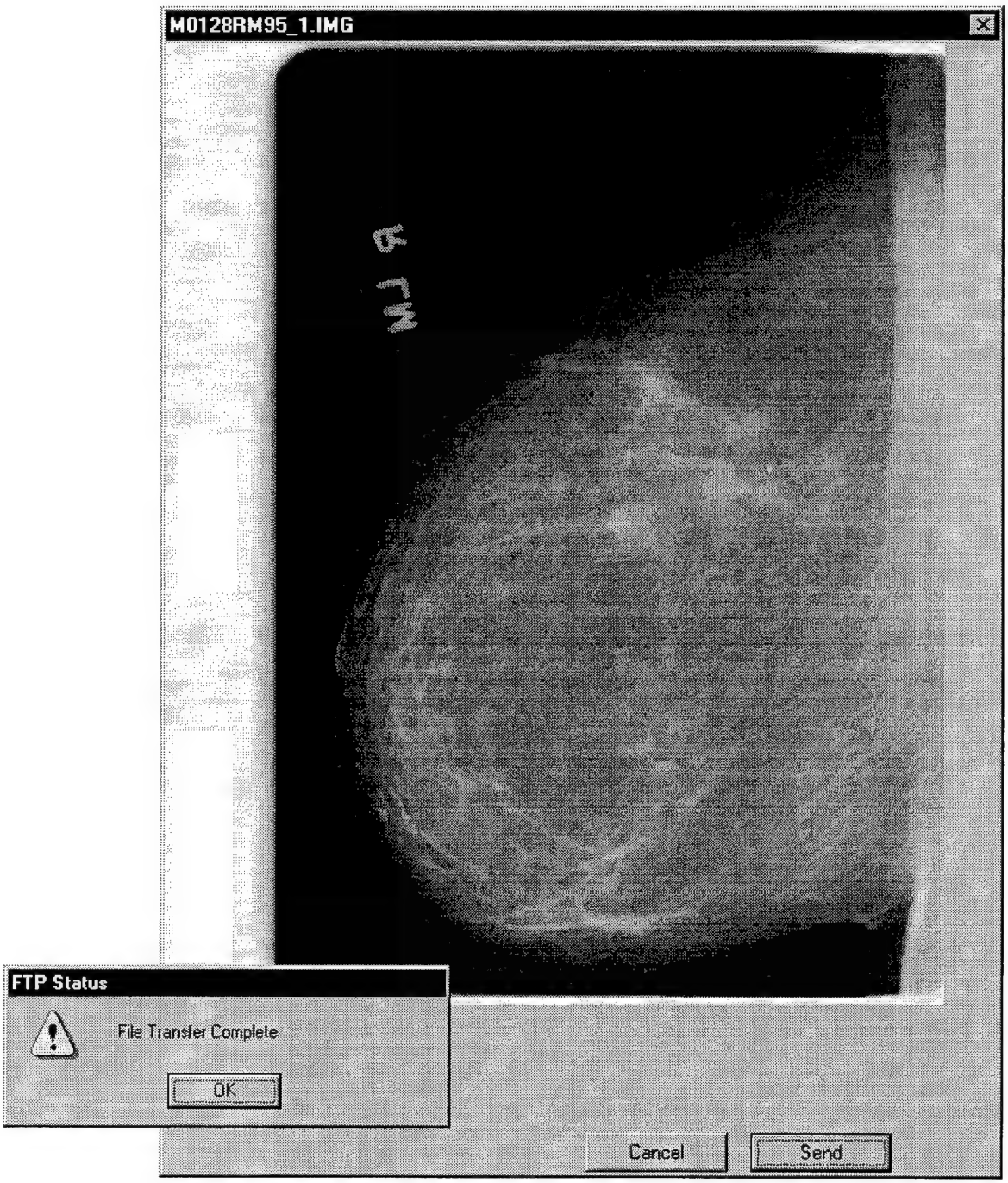


Figure 4. Digitized mammogram which is already transferred to the Jukebox.

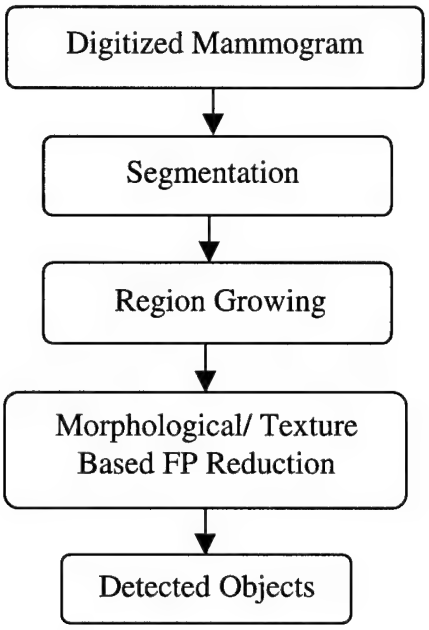


Figure 5. Mass detection algorithm.

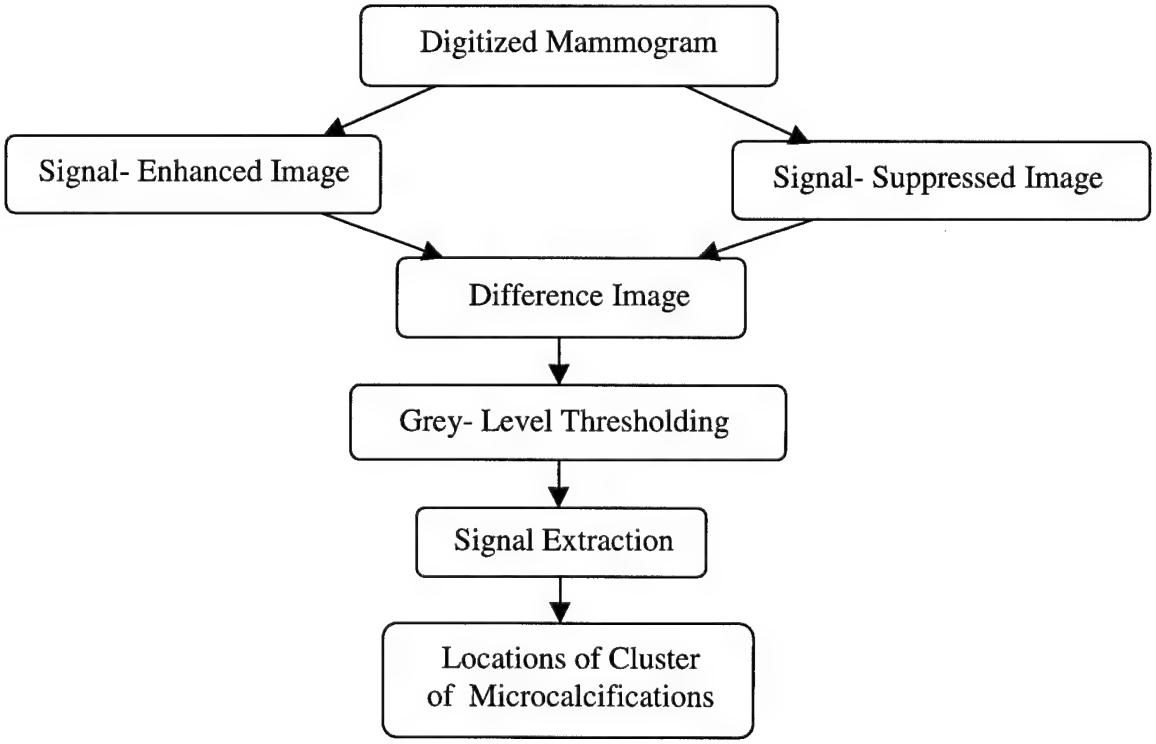
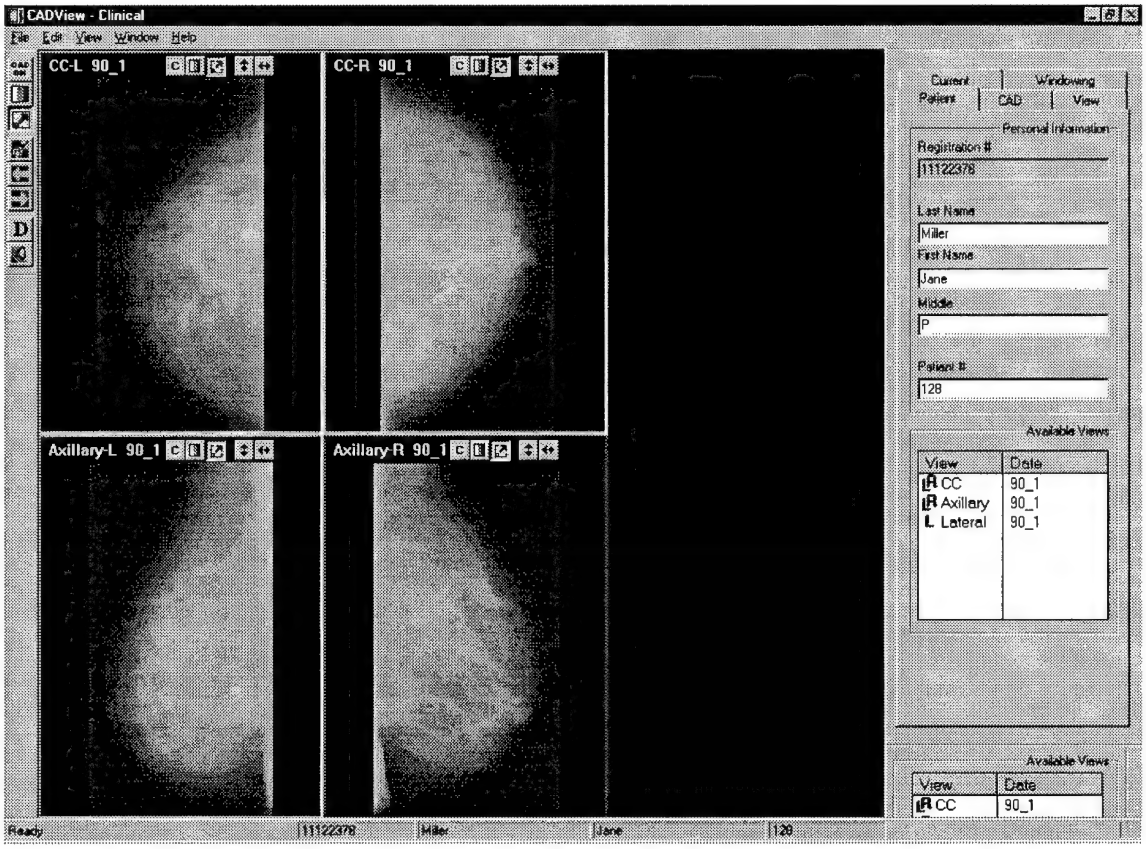
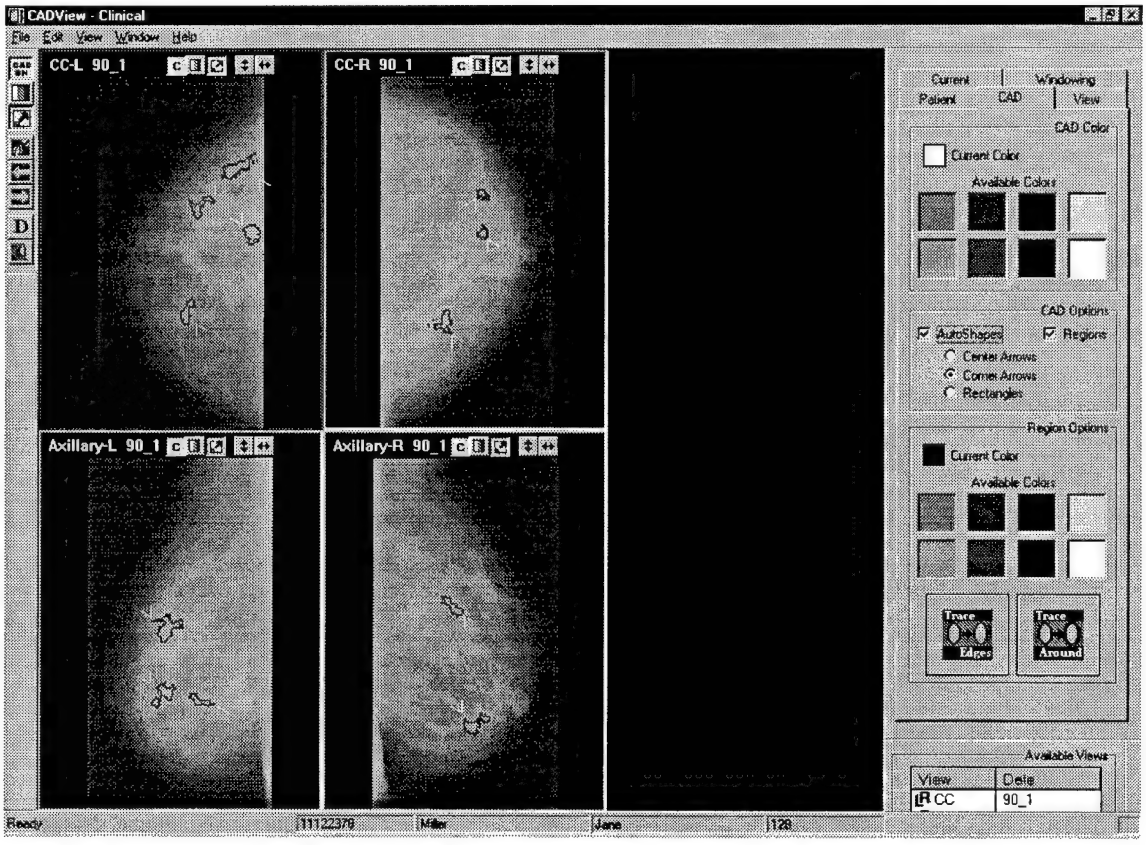


Figure 6. Microcalcifications detection algorithm

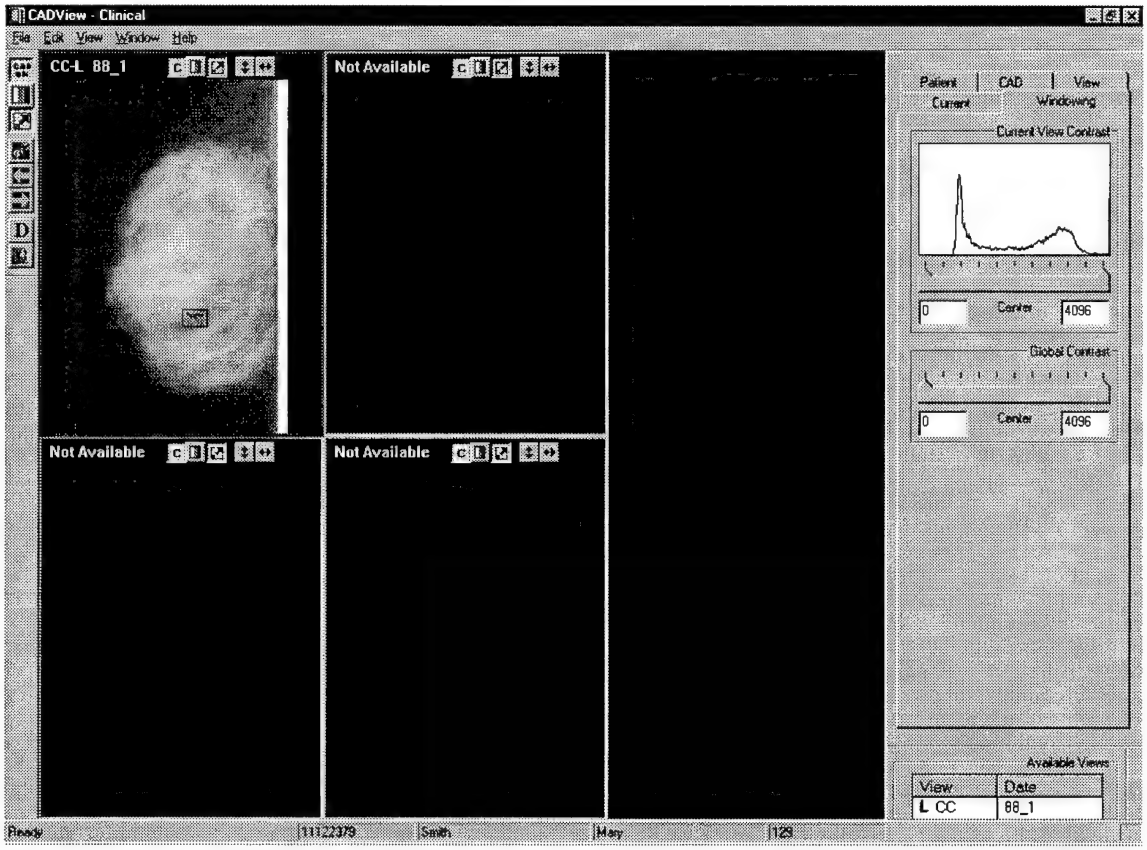


a

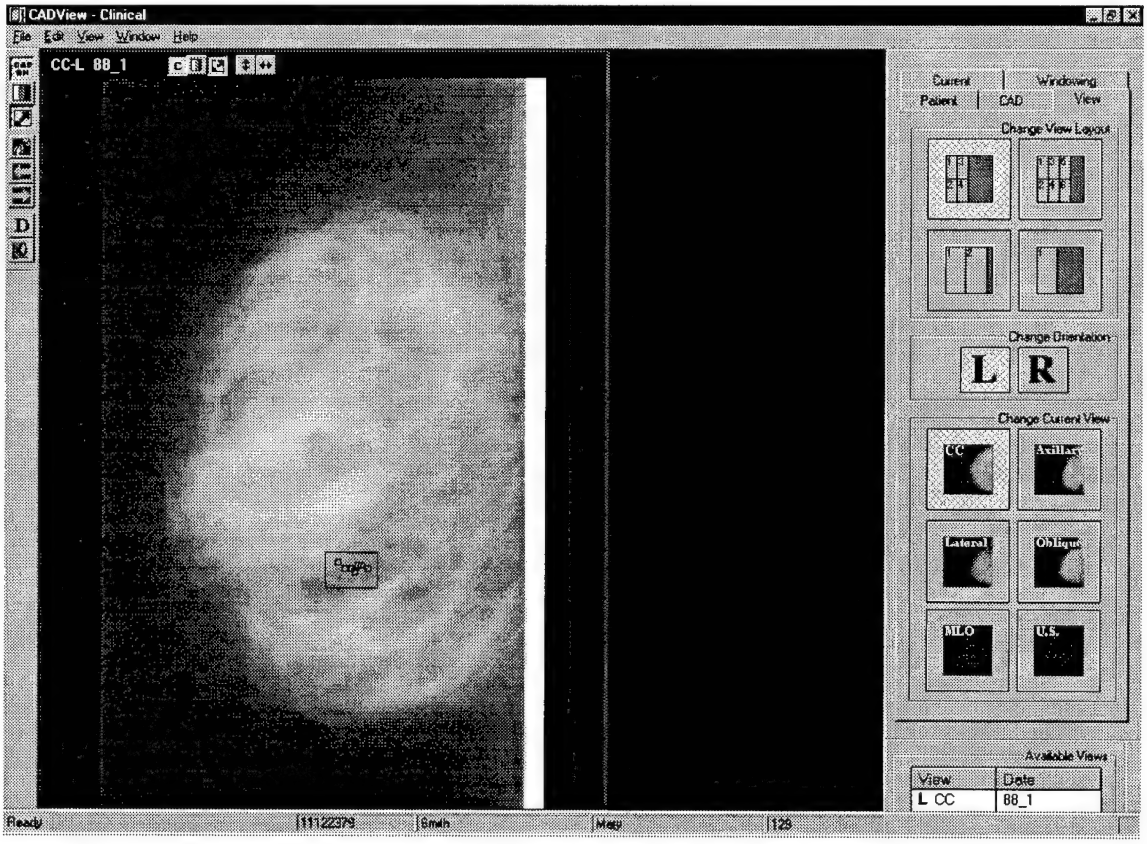


b

Figure 7. Visualization user interface. (a) Display of the patient information and CC and Axilliary views (left and write) of the digitized mammograms. (b) Display of the CAD Setup menu and CAD results as arrows pointing the masses and outlines of the masses.



a



b

Figure 8. Visualization user interface. (a) Display of the local and global image windowing menus as well as CC view with microcalcification CAD results. (b) Display of the window configuration setup menu and magnified image with microcalcification CAD results.

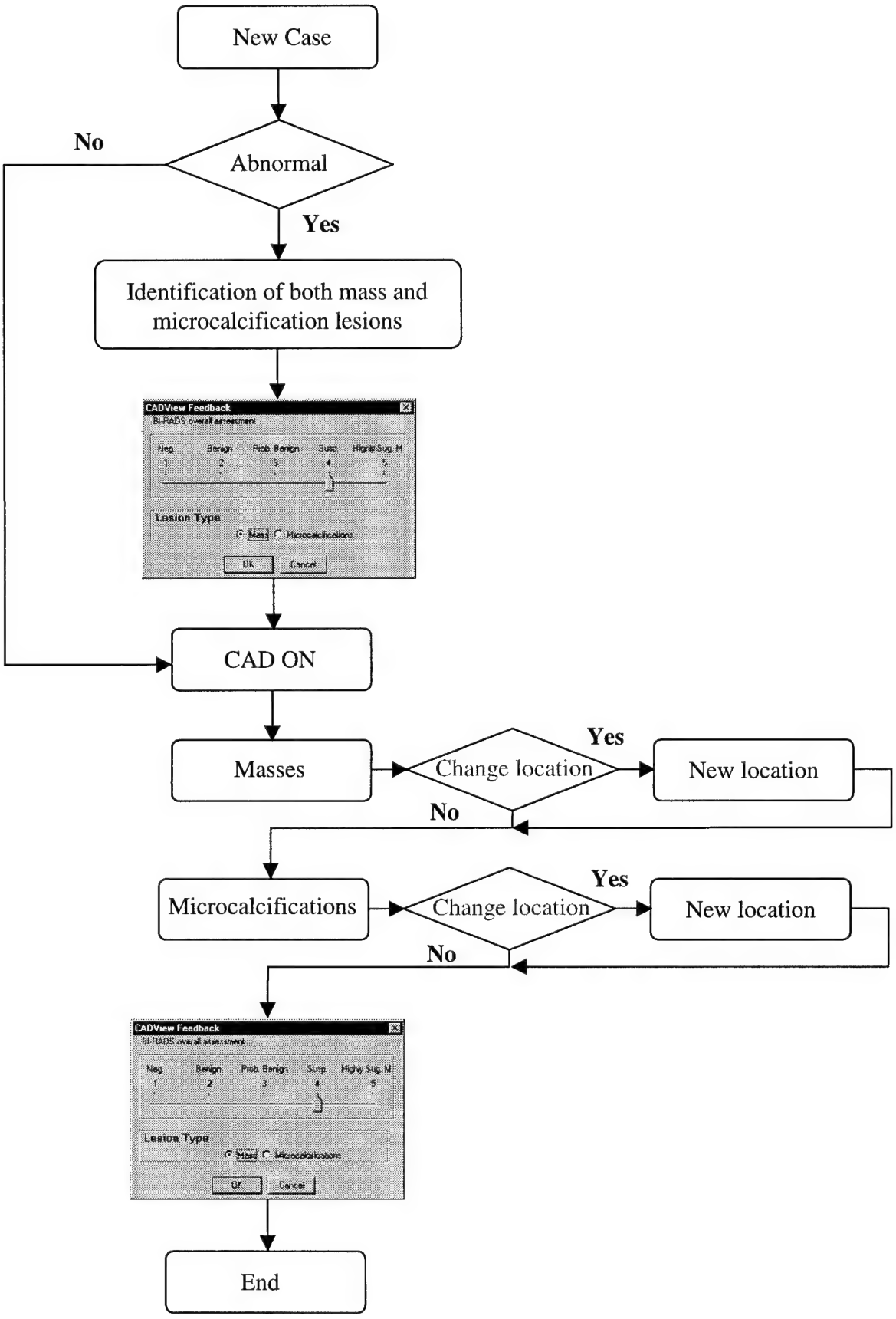
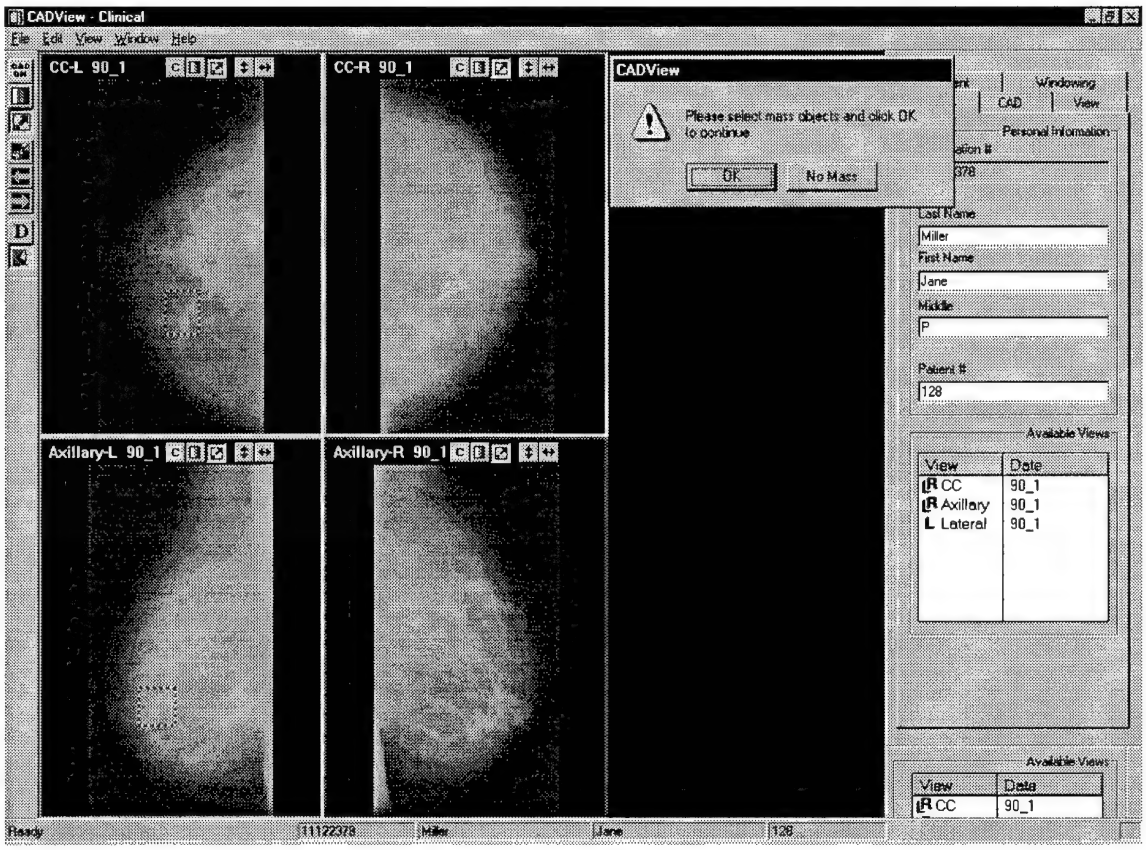
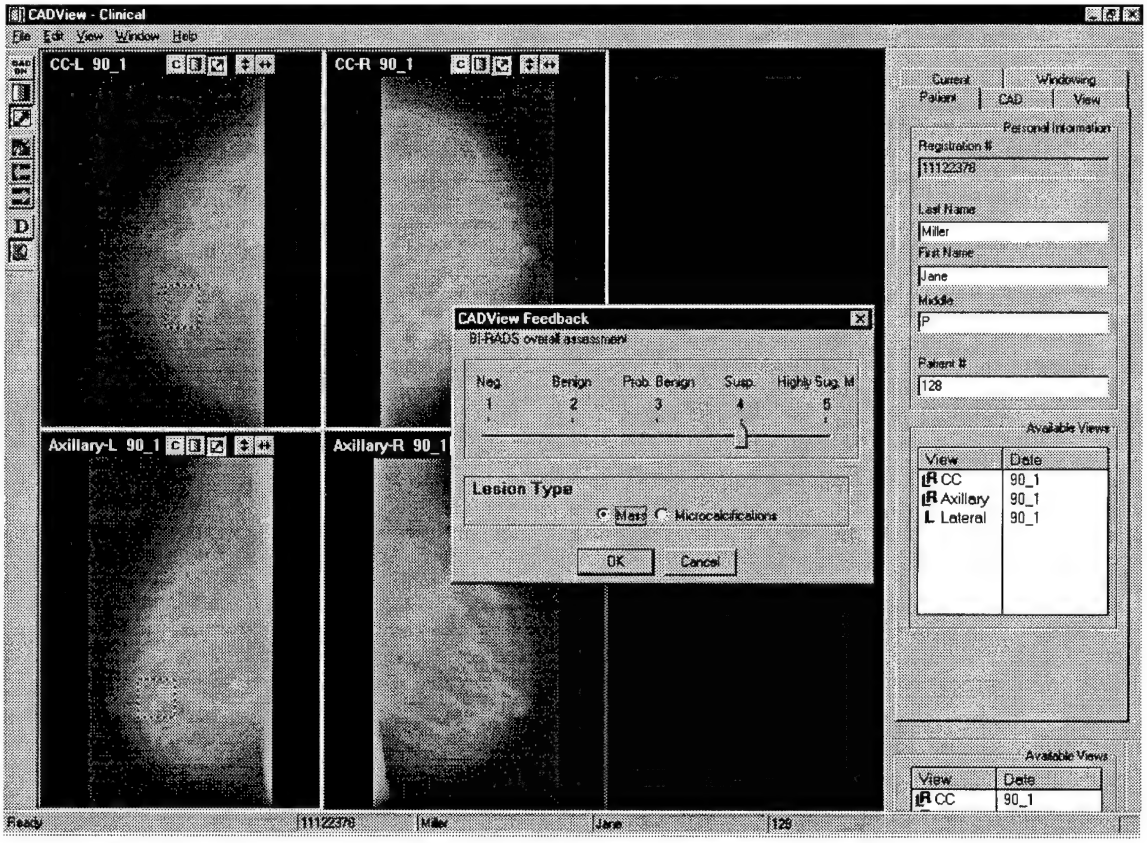


Figure 9. Sequence of the feedback information collection from the radiologist.

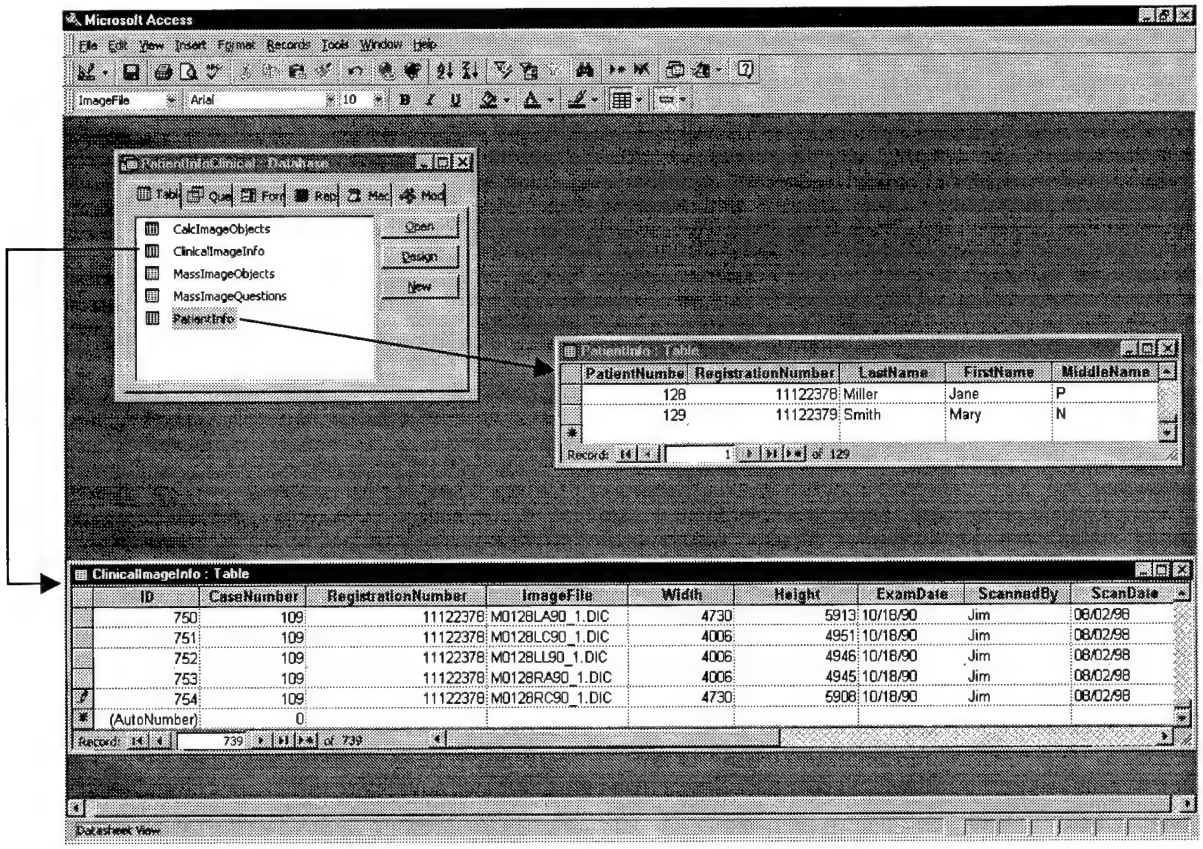


a

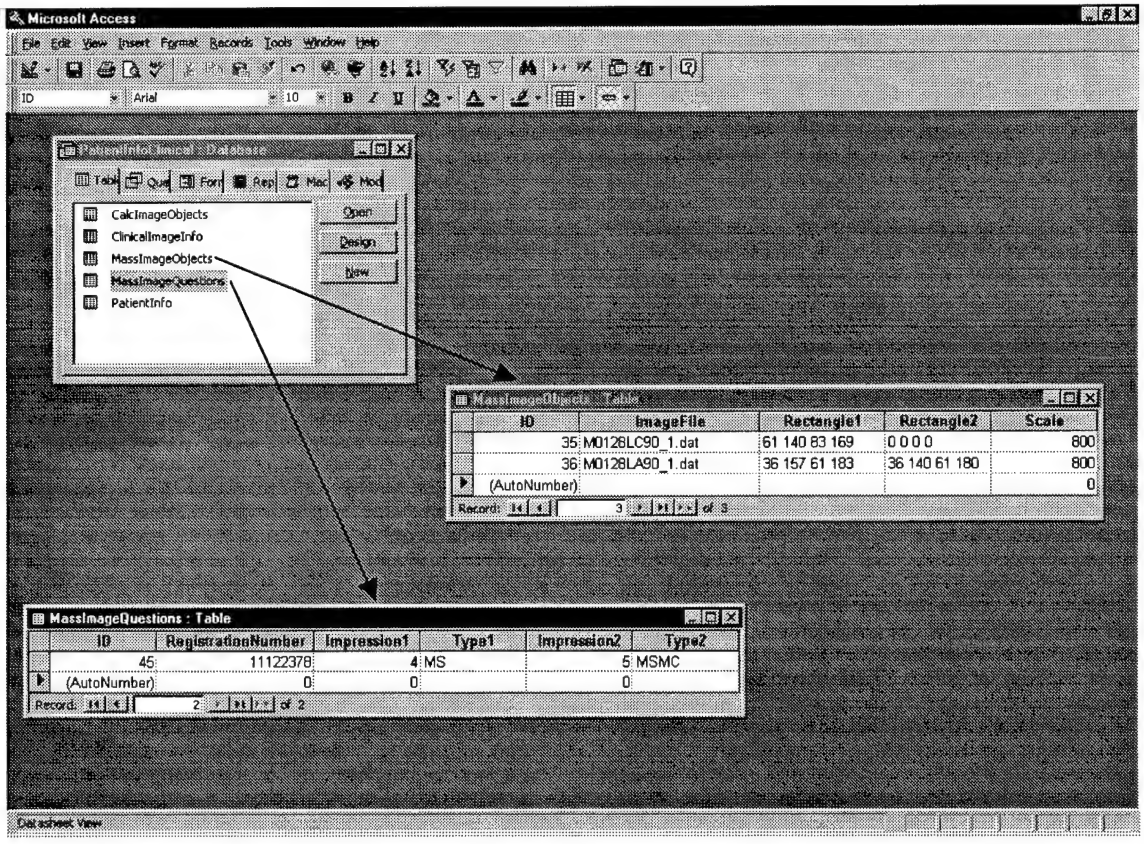


b

Figure 10. Radiologist feedback user interface. (a) Display of the radiologist marked true mass locations. (b) Display of the window with the questions to the radiologists.



a



b

Figure 11. Database of the CAD clinical system. (a) Patient information and digitized mammograms database tables. (b) Radiologist feedback database tables for the marked objects and answered questions.

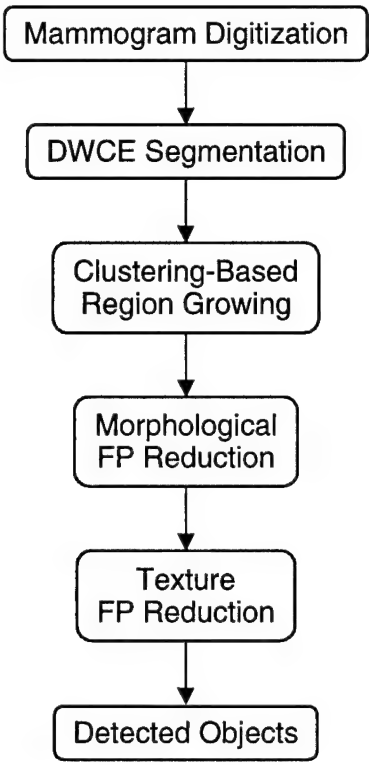


Fig. 12. Block diagram of the current mass segmentation method.

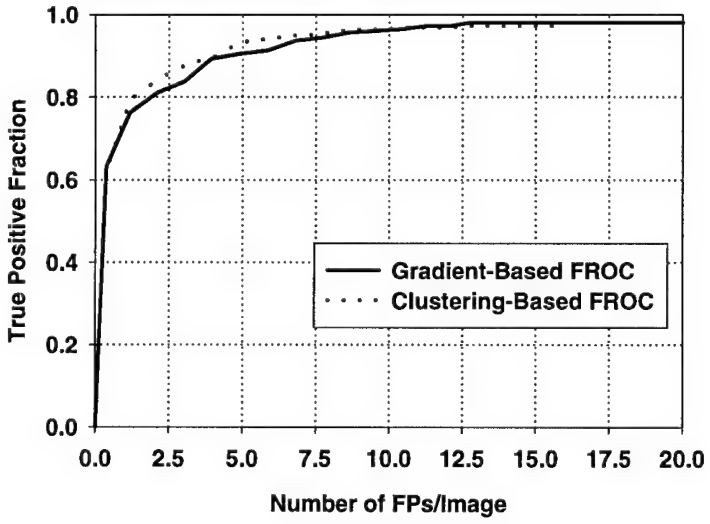


Fig. 13. The overall performance achieved by the gradient-based and clustering-based segmentation schemes.

Georgetown University

The researchers at the Georgetown University have been evaluating microcalcification and mass detection algorithms and investigating new image compression methods for mammograms. The following summaries their progresses.

(a) Preliminary clinical study using CAD system for the detection of microcalcifications

The research team at the Georgetown University has conducted a preliminary clinical study with the CAD system for the detection of microcalcifications. In this prospective clinical study, the radiologists used hard-copy printed out from the CAD program to perform the clinical trial. Since false positive detections of microcalcifications by CAD systems are a distraction to the radiologist and raise questions as to the eventual clinical utility of CAD systems. We have carefully analyzed the mammographic findings that appear in the locations of CAD detections and have counted and classified them.

a.1. Experimental methods

Two different series were run representing two different settings of the CAD algorithm. In Series 1, 200 mammogram images were analyzed. In Series 2, 95 mammogram images were analyzed. The settings for the algorithm were changed between Series 1 and 2. In Series 1, the parameters were set to detect a minimum of three suspected microcalcification foci with an average convolution neural network (CNN) output value of 0.7. In Series 2, we set the algorithm to detect a minimum of four suspected foci of microcalcification with an average CNN output value of 0.8 as the threshold.

Abnormalities seen at the sites of CAD localization were classified as representing artifacts, true positive findings and false negative findings. We included normal non-calcified punctate anatomic structures as artifacts in this analysis.

The CAD program was run prior to the mammograms being interpreted by the radiologist. Cases were selected from the clinical cases of the breast cancer screening service. Case selection required that each patient has both a current and a prior study and to have images of both breasts. Once these criteria were met, the cases were assigned or not assigned to the CAD group by selecting every other case. Cases were digitized at 100 microns using a Lumiscan (model 150) film scanner. They were then processed by the CAD program and the results returned later that day to the radiologist for assessment. The radiologist, who by then had interpreted the mammograms for the official clinical report, proceeded to review the CAD findings and classify any identified abnormalities based on examination of the original mammography film with a 2X or a 5X magnifying lens. Only one indeterminate cluster of microcalcifications was detected by the CAD program that had not been detected by (in this case) either of the two radiologists who had interpreted the study. The cluster was stable and had been missed by both the radiologist initially interpreting the older study and the radiologist interpreting the newer study. Because it was stable, no additional evaluation was done of this cluster.

In many of the sites identified by the CAD algorithm, there was more than one finding that could have resulted in the CAD detection. We chose to code these findings separately. Because of this, there are many more false positive detections indicated than the number of false positives per image would suggest. We cannot assess the effect of multiple artifacts or combinations of true microcalcifications and artifacts in the performance of the CAD program, and so we chose to record all findings. In assessing the number of false positive detections per image, we looked at each site that was recorded. If

at least one microcalcification was present along with the non-calcium structures, we graded that as a true detection for calcifications in determining the number of false positives per image. We did separate calculations for the number of false positives per image using the criteria of 1 or more and 2 or more microcalcifications in the identified field.

a.2. Findings in true negatives, true positives, false negatives, and false positives

a.2.1. True negatives

In Series 1, 44% of the mammogram films had no CAD detections and no clusters of calcifications were seen when the radiologist re-assessed the film. In Series 2, 31% of the mammogram films had no CAD detections and no clusters of calcifications on film re-assessment. **These results indicated that our system does not produce false positive on every mammogram.**

a.2.2. True positives

True positives as defined in this study, were detections with one or more small benign calcifications or indeterminate microcalcifications. In this series, the true positive detection rate was 86% in Series 1 and 94% in Series 2 when measured against a single radiologist's interpretation of the mammographic images with the CAD output and when using the presence of at least one microcalcification as a true positive detection. Overall, because we recorded separately each finding in a location identified by the CAD program, 29% of the details found in regions identified by the CAD program in Series 1 and 27 % of the detections in Series two were true positives. Vascular calcifications were considered to be false positives.

When tested previously with a proven set of cases, the CAD algorithm performance was 87% true positive detection rate at 0.5 false positive clusters per image. (Lo 1995).

a.2.3. True positive and true negative findings combined

If one combines the true negatives and true positive cases, 73% of the mammogram films in Series 1 and 58% of the films in Series 2 were correctly classified.

a.2.4. False negatives

False negative detections were defined as cases in which a benign or indeterminate cluster of microcalcifications were present on the mammogram film, but was not detected by the CAD algorithm. False negative results were seen in 8% of films in Series 1 and 3% of films in Series 2.

a.2.5. False positive detections

False positive detections accounted for 71% of the details recorded in Series 1 and 73% of the details in Series 2. As previously stated a false positive location could have multiple details within it that could explain the detection and each was recorded separately.

a.2.6. False positive detections per image

In recording the number of artifacts, we recorded separately each of the types of artifacts found in any CAD defined area of abnormality. In assessing the number of false positives per image, we accepted any CAD identified location as being a true positive if one or more true microcalcifications were present at the same site. A false positive was a location indicated by the CAD program in which calcifications were not present. Using these criteria, in Series 1 there were an average of 0.7 false positive detections per image and in Series 2 there were 0.9 false positive detections per image. However, more than one-third of images were correctly identified as normal without showing false positive as described in section 1.2.1.

If one uses the criteria of two or more calcifications for a true positive detection, in Series 1, the false positive rate was 0.8 per image and for Series 2, the rate was 1.0 false positive detections per image. See (Freedman 1997) for detailed analysis of false positive detections.

a.3. Conclusion of the preliminary studies

False positive detections in computer aided microcalcification programs are not random responses of the computer algorithm to unknown features. Better understanding of their causes should promote algorithm modification. Since the computer algorithm is, in general, responding to true punctate or short linear findings that resemble microcalcifications, this suggests that computer aided systems will function best with high-quality artifact-free films and that computer detection systems may need to be combined with improved classification systems to decrease the number of false positive detections.

(b) Mass detection using sector features with a multiple circular path neural network

In this study, our goal was to extract clinically suspicious lesions. The study was conducted with the following steps: (1) use background correction method and morphological operations to extract radio-opaque areas, (2) delineate the boundary of the areas, (3) compute the features and texture of the masses with emphasis on the boundary, (4) design and plan training strategy using a neural network as classifier for the recognition of mass features. An overall detection scheme of our proposed framework is shown in Figure 1.

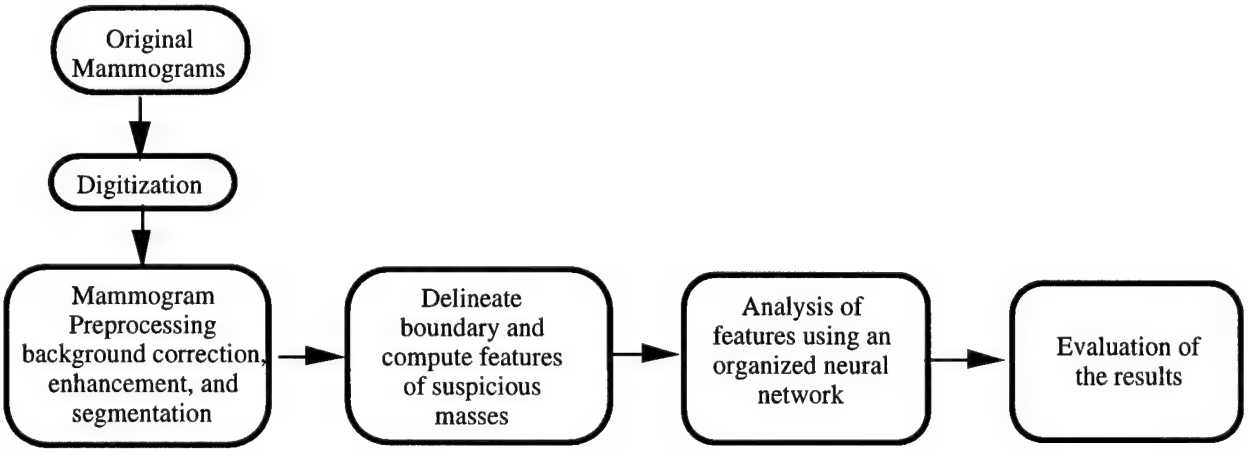


Figure 1. A flowchart for the detection of masses in this study.

b.1. Morphology-based preprocessing for image consistency and mass enhancement

Mathematical morphology is powerful in analyzing and describing geometrical relations and is a formalization of intuitive concepts such as size or shape. The two basic morphological operations are "erosion" and "dilation," which are consistently defined for binary and gray-scale images. Using these two basic operations, two other basic and important operators, "opening" and "closing", can be defined as follows:

$$\text{opening:} \qquad \mathbf{X}_{\mathbf{B}} \equiv (\mathbf{X} \ominus \boldsymbol{B}) \oplus \boldsymbol{B}, \qquad \text{...(1)}$$

$$\text{closing:} \qquad \boldsymbol{X}^{\boldsymbol{B}} \equiv (\boldsymbol{X} \oplus \boldsymbol{B}) \ominus \boldsymbol{B}, \qquad \text{...(2)}$$

where $\boldsymbol{X}$ indicates the original image, $\boldsymbol{B}$ represents the structuring element, and $\oplus$ and $\ominus$ indicate the operations "dilation" and "erosion," respectively. Based on the "opening" operation, we have developed an operation for background correction. The operation is represented by

$$\boldsymbol{X} - \boldsymbol{X}_{\boldsymbol{B}} = \boldsymbol{X} - (\boldsymbol{X} \ominus \boldsymbol{B}) \oplus \boldsymbol{B}. \qquad \text{...(3)}$$

This equation represents the subtraction of the image processed by the operator "opening" from the original image.

Figure 2 shows the effect of the operation represented by Eq. (3): (a) illustrates a structuring element, (b) shows the original signal (gray line) and the processed signal (black line) by "opening", and (c) denotes the final output signal of the operation indicated by Eq. (3). (c) is the subtraction of the black line signal from the gray line signal in (b). Note that the detected peak signals were not affected by the operation. Hence the mass signals detected by the operation retain their original shapes.

As can be seen in this graph, the size of the detected peak significantly depends on the size of the structuring element. All peaks, which are smaller than the structuring element, can be detected. In our mass detection process, a 52 pixel-diameter structuring element will be used to detect masses whose sizes are less than 52 pixels in diameter. An object with a diameter of 52 pixels in a 512×625 pixel reduced image occupies 250 pixels in its original digitized image, and its real size is expected to be about 2.5 cm.

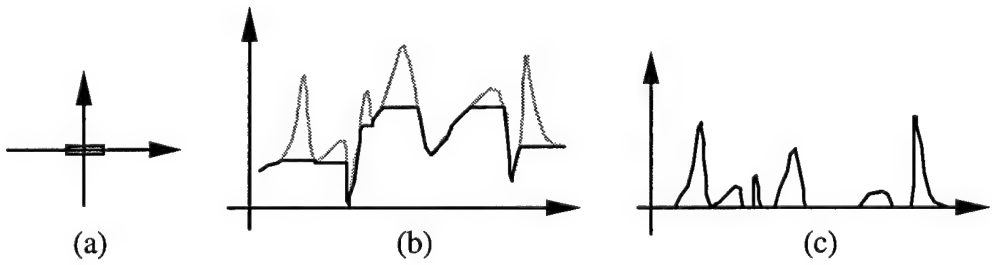


Figure 2. Effect of operation in Eq. (3): (a) structure element, (b) original signal (gray line) and signal after opening (black line), and (c) output signal of operation in Eq. (3).

b.2. Feature extraction of masses

We performed boundary detection algorithm on suspected masses which were extracted on the morphologically enhanced mammograms. A region growing with valley blocking technique was employed to delineate all the suspected areas. Then, the boundary was divided into 36 sectors (i.e., 10^{o} per sector) using 36 equi-angle dividers radiated from the center of suspicious area. The following features were computed within each 10^{o} sector of the area:

(a) "l" - the length from the center of mass to the shortest boundary segment.

(b) "a" - the normal angle of the boundary segment (or the value of cos(a)).
(c) "g" - the average gradient of gray value on the segment along the radial direction.
Technically speaking, this set of gradient values may also serve as a fuzzy system for the input layer in the neural network to be described.
(d) "c" - the gray value difference (i.e., contrast) along the radial direction.
(average gray value (h_i) calculated from the mass area located at "l"/3 inside the boundary and the average background value (b_o) calculated from the peripheral area near "l"/3 outside of the suspicious area).

Hence, a total of 144 computed features (4 features/sector for 36 sectors) can be used as input values for the analysis of suspicious areas. The relationship between the computed features and BI-RADS descriptors are discussed below:

(1) Mass Size -
The 36 "l" values would provide sufficient data for the neural network to determine the size.
(2) Mass Shape (round, oval, lobulated, or irregular) -
The 36 "l" and 36 "a" values could approximate the shape of a mass.
(3) Mass Margin (circumscribed, microlobulated, obscured, ill-defined, or spiculated) -
The 36 "g" and 36 "l" values should be able to describe the characteristics of the mass margin.
(4) Mass Density (fat-containing, low density, isodense, or highly dense) -
The 36 "c" and 36 "g" values would be able to describe the density of the mass.

In short, the BI-RADS descriptors were used as primary consideration in our feature selection. The reason for using 36 values for each nominated feature is four-fold: (a) mass boundary varies, it is difficult to describe an image pattern using a single value; (b) due to the general shape of the masses, the features of masses can be easily analyzed by the polar coordinate system; (c) in case some features are inaccurately computed in several directions due to the structure noises, such as the breast slender lines, there may still exist a sufficient number of correct features; (d) generally more accurate results can be produced by using subdivided parameters rather than using global parameters in a pattern recognition task. Other computational features (e.g., difference entropy (Li 1997) and other higher order features) are eligible but require further investigation.

b.3. A neural network system specifically designed for the extracted boundary features

We have developed a multiple circular path neural network (MCPNN) to instruct the neural network in analyzing sector features. Basically, we designed several neural network connections between the input and the first hidden layers as shown in Figure 3. Figure 3(a), (b), and (c) illustrate the full connection, a self correlation (SC) networking, and a neighborhood correlation (NC) networking, respectively. Note that the input and hidden nodes should be completely matched when combining more than one path in the study. In this case, the correlation layers only function as branch connections between input and hidden layers. When using NC paths, networking engagement within multiple sectors (e.g., 20^o, 30^o, 40^o, and 50^o of the neighborhood correlation) can be grouped.

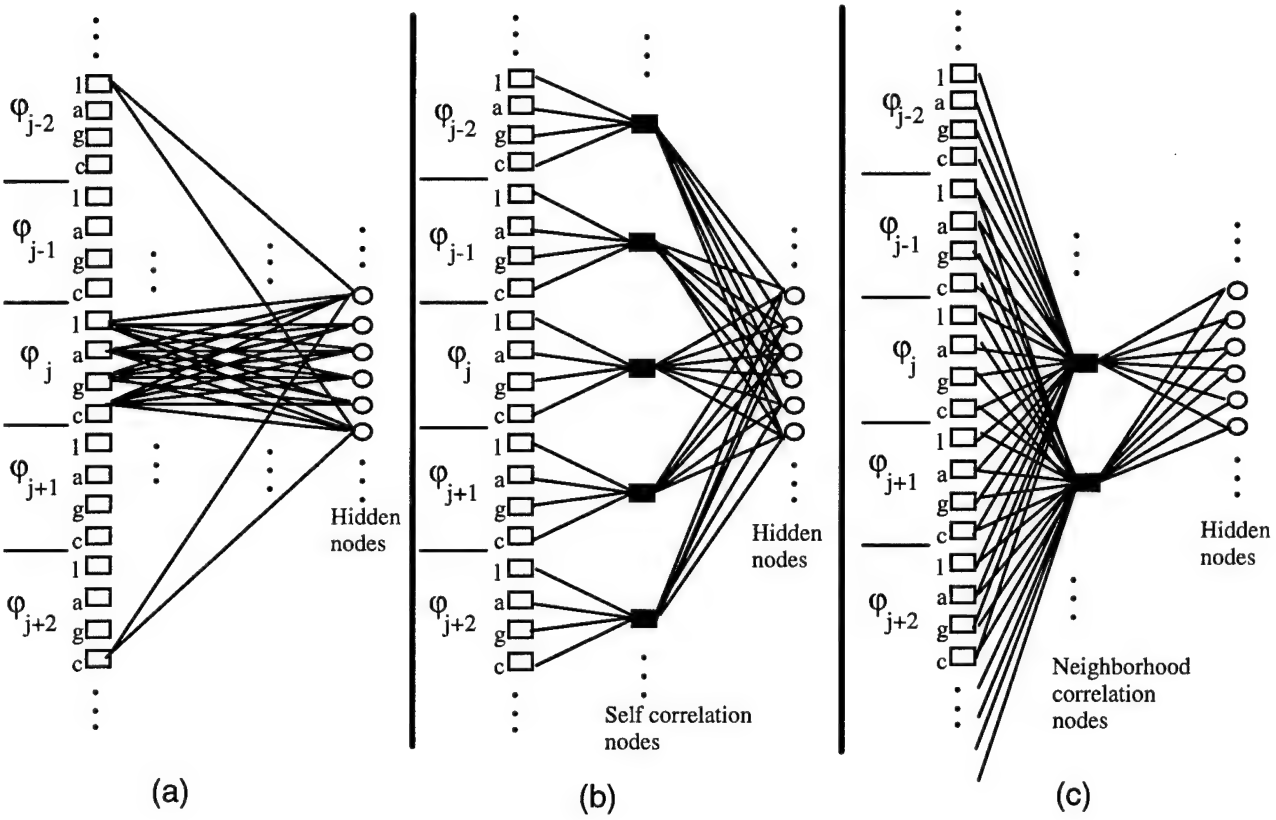


Figure 3. Three types of network paths connecting the input and the hidden layers:
(a) Full connection.
(b) A self correlation (SC) path; each node on the layer connects to a single set of the features (l,a,g,c) for the fan-in and fully connects to the hidden nodes for fan-out.
(c) A neighborhood correlation (NC) path; each node on the layer connects to five adjacent sets of the features for the fan-in and fully connects to the hidden nodes for fan-out.
Note that the fan-in nets emphasizing self correlation in (b) and neighborhood correlation in (c) represent convolution weights (i.e., the same type of sectors possess the same set of weighting factors).

b.4. Summary of feature extraction methods and the MCPNN

We have described our approach on the feature extraction, the design of MCPNN, and its corresponding training method. Figure 4 shows a flow diagram of the proposed method. Since the MCP only alters the input data connection from the input to the first hidden layer, any learning algorithm can be applied within the neural network. For simplicity, we used the back propagation algorithm for both the conventional and proposed neural network systems in the following experiments.

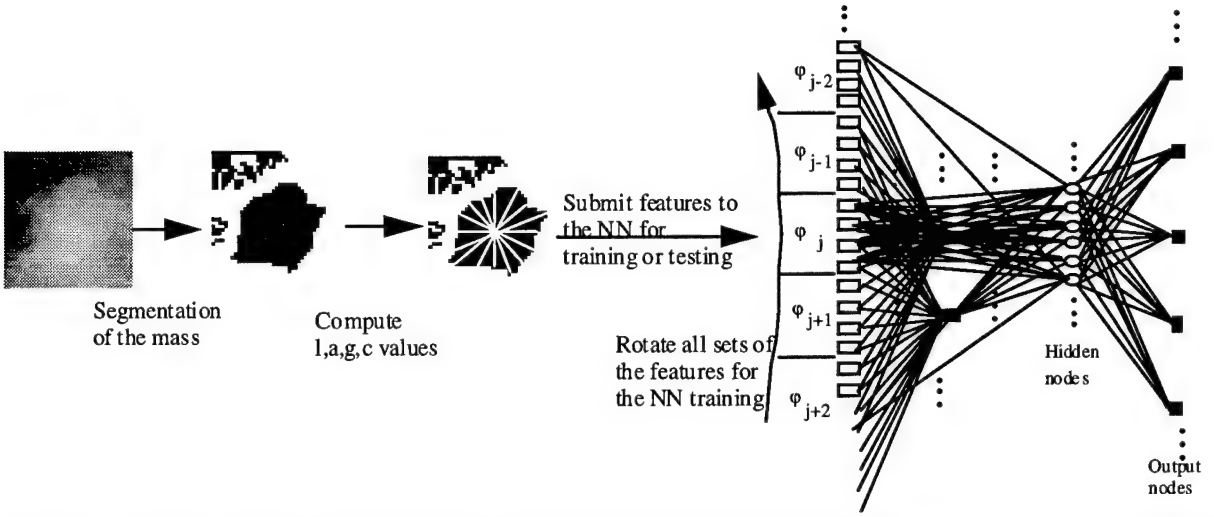


Figure 4. A flow chart, involving the MCPNN and sector features of masses, was used in the following study.

b.5. Experiments and results

We selected 91 mammograms and digitized each mammogram with a computer format of 2048×2500×12 bits (for an 8"×11" area where each image pixel represents 100 μm square). No two mammograms were selected from the same patient film jacket. All the digitized mammograms were miniaturized to 512×625×12 bits using 4×4 pixel averaging and were processed by the above methods to perform mass detection. Based on the corresponding biopsy reports, one experienced radiologist read all 91 mammograms and identified 75 areas containing masses. (Note that the reports recorded the malignancy of the biopsy specimens. The radiologist only used them as reference for the identification of masses.) Through the pre-process and the first step screen based on the circularity test, a total of 125 suspicious areas were extracted from the 91 digitized mammograms.

Experiment 1

We randomly selected 54 computer-segmented areas where 30 patches were matched with the radiologist's identification and 24 were not. This database was used to train two neural network systems: (1) a conventional 3-layer BP neural network (with 125 nodes in the hidden layer) and (2) the proposed MCP training method using the same neural network learning algorithm. The structure of the MCPNN was described earlier. However, we used one fully connected path, four SC paths, four 20° NC paths, four 30° NC paths, three 40° NC paths, and two 50° NC paths in the first step network connection for the MCPNN. Both neural network systems were trained by the error back propagation algorithm by feeding the features from the input layer and registering the corresponding target value at the output side. Once the training of the neural networks was complete, we then used the remaining 71 computer segmented areas for the testing. None of the images and their corresponding patients in the testing set could be found in the training set. The neural network output values were fed into the LABROC program (Metz 1989) for the performance evaluation. The results indicated that the areas (Az) under the receiver operator characteristic (ROC) curves were 0.781 and 0.844 using the conventional BPNN and the MCPNN, respectively. The ROC curves of these two neural network training methods are shown in Figure 5(A). We also invited another senior mammographer to conduct an ROC observer study. The mammographer was asked to rate each patch using a numerical scale ranging 0-10 for its likelihood of being a mass. These 71 numbers were also fed into the LABROC program. The mammographer's performance in Az on this set of test cases was 0.909. The corresponding ROC curve is also shown in Figure 5(A).

Experiment 2

We also conducted a leave-one-case-out experiment using the same database. In this experiment, we used those patches extracted from 90 mammograms for the training and used the patches (most of them are single) extracted from the remaining one mammogram as test objects. The procedure was repeated 91 times to allow every suspicious patch from each mammogram to be tested in the experiment. For each individual suspicious area, the computed features were identical to those used in Experiment 1. Again, both neural network systems were independently evaluated with the same procedure. The results indicated that the Az values were 0.799 and 0.887 using the conventional back propagation neural network and the MCPNN, respectively. Figure 5(B) shows the ROC curves of these two neural network systems using the leave-one-of-out procedure in the experiment.

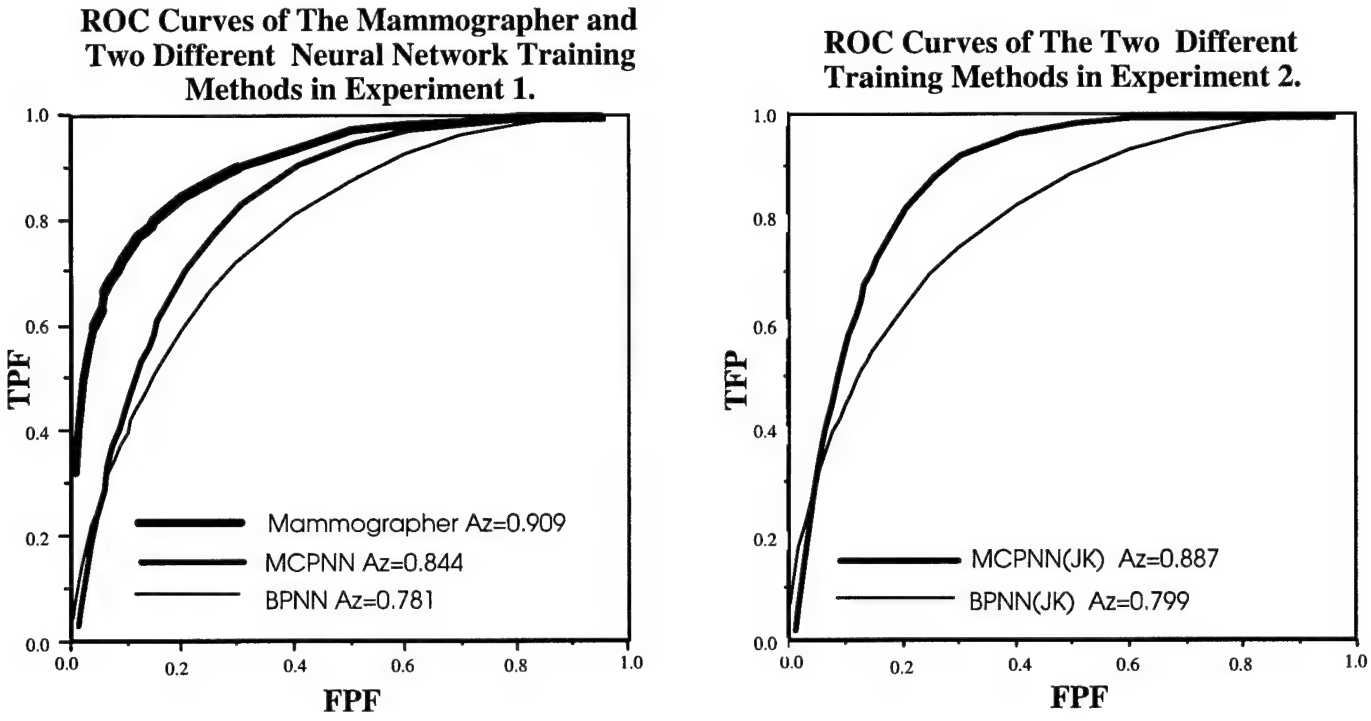


Figure 5. The ROC curves obtained from corresponding experiments.

(A) The left figure shows that the performance of MCPNN training method is superior to that of the conventional input method. The highest curve is the ROC performance of the senior mammographer.

(B) The right figure shows similar results with a higher performance using the leave-one-case-out procedure as described in Experiment 2.

Through this study, we found that the selected features are somewhat effective in the detection of masses. In Experiment 1, we found that the performances of both neural network systems were increased. This might be due to the increased number of cases (from 54 to 124) in the training set. In Experiment 2, the Az value was improved by 0.043 using the MCPNN training method that was higher than Az difference of 0.018 obtained by the conventional training method. The results implied that the MCPNN learned more effectively than the conventional BP when the number of training cases was increased.

It is known in the field of artificial intelligence that the key factors in pattern recognition are: (1) effective methods in the extraction of features and (2) analytic methods (e.g., back propagation neural

network) for the extracted features. In this study, we showed that the training method designed to guide the analyzer is also an important factor to a successful pattern recognition task. Though this finding is not new, the trend of developing training methods for various pattern recognition tasks was not established in the field of pattern recognition. In this work, we demonstrated that organized features with proper network connection and task-oriented guidance would assist the neural network in performing the task.

Since the mass can be overlapped with glandular tissues, a significant part of the mass may be obscured and is unrecoverable by digital image processing techniques. By reviewing those failure cases, we found that substantial false-negative cases were in this category. However, these cases were correctly identified by the radiologists. This implies that we need to find a way to train the neural network to recognize those cases with sufficient sectors showing signs of masses.

(c) Integer Wavelet Computation for digital (digitized) Mammography

As described in 1997 report, we have developed a research software package that is capable of compression any image with a lossless or lossy result controllable by the user. This research package allows the user to select desired wavelet. Any wavelet can be approximated by its associated integer implementation (Lo 1997). Both lossless and lossy compression were studied on 10 sets of mammograms (4 images per set: MLO and CC views with left and right mammograms as a set) which were digitized with 100 μm and 12 bits per pixel (i.e., 11.5 Mbytes per image originally). We chose 5 large and 5 small breasts for each set in our study.

c.1. Breast area lossless compression study

We have developed a boundary detection program which can extract breast area from the original digitized mammograms. We then performed a lossless compression study using Daubechies' wavelet transform with integer computation method. The results are shown in Table I.

Table I. Compression ratios based on Daubechies' wavelet transform with integer computation.

Breast Type	Large Breast		Small Breast	
View	CC	MLO	CC	MLO
Orginal Data	2,048x2,812x2bytes (11,517,952 bytes)		2,048x2,812x2bytes (11,517,952 bytes)	
Compressed Size using Lossless	1,878,948 byte	1,972,252 byte	562,240 byte	694,690 byte
Compression based on breast	3.36	3.21	3.38	3.19
Compression based on original	6.13	5.84	20.48	16.58

We found that the average compression ratio was about 3.3 based on the original data confining in the breast area with 12 bits/pixel. This result was increased to 6 for large breast images and to about 18 for small breast mammograms.

c.2. CAD guided compression study
(lossy compression with lossless results on CAD indicated patches)

One important feature of the integer wavelet compression method is that the lossless and lossy methods can be implemented with the same transformed domain. The transformed domain coefficients used in the above lossless compression study were further processed by a quantization procedure (i.e., the values were uniformly divided by a factor in each level of wavelet compartments) followed by an arithmetic coding. At this point, we used the CAD program to guide the compression procedure to perform the lossless encoding (with non-quantization coefficients of the patches) for the suspected microcalcifications. Hence, we obtained a combined decompression result in which no error would be generated on the CAD indicated patches. The decompressed images were compared to the original images to obtain the normalized mean-square-errors (NMSE). Normalization factor is the area of the breast instead of the entire mammogram. The average results are shown in Table II.

Table II. Compression ratios and NMSE values based on Daubechies' wavelet transform with integer computation

Breast Type	Large Breast		Small Breast	
View	CC	MLO	CC	MLO
Orginal Data Size	2,048x2,812x2bytes (11,517,952 bytes)		2,048x2,812x2bytes (11,517,952 bytes)	
Compression ratio based on original image (wavelet lossy compression)	50:1	50:1	50:1	50:1
NMSE (breast only) (wavelet lossy compression)	196	215	79	86
No. of suspected patches detected by the CAD program	365	421	137	154
Compression ratio based on original image (CAD guided compression)	14:1	13:1	35:1	33:1
NMSE (breast only) (CAD guided compression)	158	176	71	79

Note that the major differences between this lossy compression study and the study reported last year are in (1) air space are completely masked and filled by a constant value in the reconstructed images and (2) the use of CAD-guided lossless compression procedure.

University of Iowa

Development of Methods for Analyzing Pilot Clinical Trial Data

We have been testing the applicability of the Dorfman, Berbaum, Metz (1992) multireader, multipatient (MRMP) methodology for analyzing receiver operating characteristic (ROC) data from the clinical trial. The CAD workstation implements the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) final categories. In clinical trials, these categories are action categories and have implications for patient care. The category "negative" translates into one year followup, "probably benign finding" translates into the course of action "short interval followup suggested," "suspicious abnormality" translates into the course of action "biopsy should be considered," and "highly suggestive of malignancy" translates into "appropriate action should be taken". Some diagnostic imaging systems may lead to more conservative or liberal actions than others. We plan to estimate decision thresholds associated with the action categories using proper ROC analysis (Dorfman, Berbaum KS, Metz et al., 1997). Proper ROC analysis is essential for this pilot clinical trial because of the paucity of cancers.

We have tested the Dorfman/Berbaum/Metz (DBM) methodology with a comprehensive series of computer simulations on factorial experimental design (Dorfman, Berbaum, Lenth et al., 1998). The results suggest that the DBM method provides trustworthy alpha levels with discrete ratings when ROC area is not too large, and case and reader sample sizes are not too small. In other situations, the test tends to be somewhat conservative or very slightly liberal. We have also tested the DBM methodology with a comprehensive series of computer simulations on split plot experimental design (Dorfman, Berbaum, Lenth et al., 1999). Our Monte Carlo simulations show that the DBM multireader methodology can be validly extended to the split plot design where readers interpret imaging studies of different patients in CAD vs no CAD conditions. Both of these validation studies used a balanced design, which is appropriate for laboratory studies, but perhaps not for clinical trials. We are currently implementing the DBM methodology for unbalanced designs in the event that different readers finish with a different numbers of imaging studies read in CAD and no CAD conditions.

(7) Conclusions

We have completed the GUI development this year. In our revised scope of work in response to the 50% budget reduction, we had planned to implement only the microcalcification detection algorithm on the CAD workstation for the pilot clinical trial. Recently, we have decided to implement both the microcalcification detection algorithm and the mass detection algorithm on the CAD workstation for the following reasons. First, the mammographic signs of both microcalcifications and masses are equally prevalent in clinical cases. If we concentrate on microcalcification detection alone, we will reduce the number of useful cases into about half. This will result in a reduction of statistical power. Second, masses are more difficult to detect on mammograms and thus computer-aided detection of masses will be more useful to radiologists. Third, if the workstation can provide CAD for both microcalcifications and masses, the pilot clinical trial will more closely simulate clinical settings and the results may be more relevant. The addition of the mass detection algorithms essentially brings the efforts of developing the CAD workstation back to the original proposed level, i.e, the level before the 50% budget reduction. Although this will cause a reduction in the time and budget available for the pilot clinical study, we believe that this is a more proper approach to evaluate the utility of CAD.

The research teams at the University of Michigan (UM) and at the Georgetown University (GU) continue to improve their CAD algorithms for detection of microcalcifications and masses. After evaluating the UM and GU algorithms separately, our next step is to combine the most effective techniques developed by the two teams into one detection program for each type of lesions. The combination of the best techniques from the two teams is expected to further improve the performance of the detection algorithms. These improved algorithms will be implemented in the CAD workstation for the pilot clinical study.

The CAD-guided image compression project is progressing as planned. The compression technique has been evaluated in a small data set described in the GU report in Section (6). After this objective evaluation of algorithm performance, we are collecting cases and will prepare laser printed films for a subjective image quality comparison study.

Because of the change in the strategy for the CAD workstation development described above, there is a delay in starting the pilot clinical study. We plan to request a no-cost-time-extension to make up for part of the work. We will submit a request to the USAMRMC Breast Cancer Research Program in a few months.

(8) References

National Center for Health Statistics. Vital statistics of the United States, 1987. Vol. 2. Mortality. Part A, DHHS Publication no. (PHS) 90-1101 (Government Printing Office, Washington, D.C., 1990).

Baines CJ, Miller AB, Wall C and al e. Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers. Radiology 1986; 160: 295-298.

Bassett LW, Bunnell DH, Jahanshahi R, Gold RH, Arndt RD and Linsman J. Breast cancer detection: one versus two views. Radiology 1987; 165: 95-97.

Bird RE, Wallace TW and Yankaskas BC. Analysis of cancers missed at screening mammography. Radiology 1992; 184: 613-617.

Boring CC, Squires TS, Tong T and Montgomery S. Cancer statistics 1994. CA-A Cancer Journal for Clinicians 1994; 44: 7-26.

Byrne C, Smart CR, Cherk C and Hartmann WH. Survival advantage differences by age: evaluation of the extended follow-up of the Breast Cancer Detection Demonstration Project. Cancer 1994; 74: 301-310.

Chan HP, Doi K, Vyborny CJ, Schmidt RA, Metz CE, Lam KL, Ogura T, Wu Y and MacMahon H. Improvement in radiologists' detection of clustered microcalcifications on mammograms. The potential of computer-aided diagnosis. Invest Radiol 1990; 25: 1102-1110.

Curpen BN, Sickles EA, Sollitto RA and al. e. The comparative value of mammographic screening for women 40-49 years old versus women 50-59 years old. AJR 1995; 164: 1099-1103.

Dorfman DD, Berbaum KS, Metz CE, Lenth RV, Hanley JA, Abu-Dagga H. Proper receiver operating characteristic analysis: the bigamma model. Acad Radiol 1997; 4:138-149.

Dorfman DD, Berbaum KS, Lenth RV, Chen Y-F, Donaghy BA. Monte Carlo validation of a multireader method for receiver operating characteristic discrete rating data: factorial experimental design. Acad Radiol 1998;5:591-602.

Dorfman DD, Berbaum KS, Lenth RV, Chen Y-F. Monte Carlo validation of a multireader method for receiver operating characteristic discrete rating data: split plot experimental design. Accepted for presentation at SPIE, February, 1999, San Diego, CA.

Feig SA and Hendrick RE, *Risk, Benefit, and Controversies in Mammographic Screening. In: Syllabus: A categorical Course in Physics Technical Aspects of Breast Imaging*, A. G. Haus and M. J. Yaffe, (Radiological Society of North America, Inc, Oak Brook, IL, 1993).

Freedman MT, Lo SCB, Artz, ST, Lau I, Mun SK, Classification of false positive findings on computer aided detection of breast microcalcifications. Proc. SPIE. 1997; 3034: 853-859.

Harris JR, Lippman ME, Veronesi U and Willett W. Breast Cancer. N Engl J Med 1992; 327: 319-328.

Harvey JA, Fajardo LL and Innis CA. Previous mammograms in patients with impalpable breast carcinomas: Retrospective vs blinded interpretation. AJR 1993; 161: 1167-1172.

Haug PJ, Tocino IM, Clayton PD and Bair TL. Automated management of screening and diagnostic mammography. Radiology 1987; 164: 747-752.

Hillman BJ, Fajardo LL, Hunter TB and al e. Mammogram interpretation by physician assistants. AJR 1987; 149: 907-911.

Kalisher L. Factors influencing false negative rates in xeromammography. Radiology 1979; 133: 297-301.

Kegelmeyer WP, Pruneda JM, Bourland PD, Hillis A, Riggs MW and Nipper ML. Computer-aided mammographic screening for spiculated lesions. Radiology 1994; 191: 331-337.

Li, H, Lo SC, Wang Y, Freedman MT, and Mun SK: Mammographic mass Detection by Stochastic Modeling and a Multi-Module Neural Network, Proc. SPIE. 1997; 3034: 480-490.

Lo SCB, Chan HP, Lin JS, et al. Artificial convolution neural network for medical image pattern recognition. Neural Networks. 1995. 8:1201-1214.

Lo SCB, Xuan J, Li H, Wang YJ, Freedman MT, and Mun SK, Dyadic decomposition: A unified perspective on predictive, subband, and wavelet transforms," Proc. SPIE. 1997; 3031: 286-301.

Martin JE, Moskowitz M and Milbrath JR. Breast cancer missed by mammography. AJR 1979; 132: 737-739.

Metz CE and Shen S. Gains in diagnostic accuracy from replicated readings of diagnostic images: prediction and assessment in terms of ROC analysis. Med Decis Making 1992; 12: 60-75.

Metz CE, Shen JH, and Herman BA: New methods for estimating a binormal ROC curve from continuously distributed test results. Presented at the 1990 Annual Meeting of the American Statistical Association, Anaheim, CA.

Moskowitz M, *Benefit and risk. In: Breast Cancer Detection: Mammography and Other Methods in Breast Imaging*, 2nd ed. Eds. L. W. Bassett and R. H. Gold, (Grune and Stratton, New York, 1987).

Seidman H, Gelb SK, Silverberg E, LaVerda N and Lubera JA. Survival experience in the Breast Cancer Detection Demonstration Project. CA Cancer J Clin. 1987; 37: 258-290.

Shtern F, Stelling C, Goldberg B and Hawkins R. Novel technologies in breast imaging: National Cancer Institute perspective. Society of Breast Imaging, Orlando, Florida, 1995; 153-156.

Smart CR, Hendrick RE, Rutledge JH and Smith RA. Benefit of mammography screening in women ages 40 to 49 years: current evidence from randomized controlled trials. Cancer 1995; 75: 1619-1626.

Tabar L and Dean PB, *Teaching Atlas of Mammography*, (Thieme, New York, 1985).

Thurfjell EL, Lernevall KA and Taube AAS. Benefit of independent double reading in a population-based mammography screening program. Radiology 1994; 191: 241-244.

Wallis MG, Walsh MT and Lee JR. A review of false negative mammography in a symptomatic population. Clinical Radiology 1991; 44: 13-15.

(9) Appendix

Publications (University of Michigan)

Journal Articles

1. Sahiner B, Chan HP, Petrick N, Helvie MA, Goodsitt MM. Development of a high-sensitivity classifier for computer-aided diagnosis: Application to classification of malignant and benign masses. Physics in Medicine and Biology 1998; 43: 2853-2871.

2. Chan HP, Sahiner B, Lam KL, Petrick N, Helvie MA, Goodsitt MM, Adler DD. Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces. Medical Physics. 1998; 25: (October, in press).

3. Chan HP, Sahiner B, Helvie MA, Petrick N, Roubidoux MA, Wilson TE, Adler DD, Paramagul C, Newman JS, Sanjay-Gopal S. Improvement of radiologists' characterization of mammographic masses by computer-aided diagnosis: an ROC Study. Radiology (Accepted pending revision).

4. Petrick N, Chan HP, Sahiner B, Helvie MA, Goodsitt MM. Combined adaptive enhancement and object-based region growing for automated detection of masses on mammograms. Medical Physics (Submitted).

Conference Proceedings

1. Chan HP, Sahner B, Wagner RF, Petrick N. Effects of sample size on classifier design for computer-aided diagnosis. Proc. SPIE 1998; 3338: 845-858.

2. Sanjay-Gopal S, Chan HP, Petrick N, Wilson T, Sahiner B, Helvie MA, Goodsitt MM. A regional mammogram registration technique for automated analysis of interval changes of breast lesions. Proc. SPIE 1998; 3338: 118-129.

Abstracts, Presentations, Scientific Exhibits

1. Sahiner B, Chan HP, Chenevert T, Petrick N, Helvie MA, Sanjay-Gopal S. Computer-aided characterization of malignant and benign lesions on breast MR images using texture features. Presented at the 83rd Scientific Assembly and Annual Meeting of the Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 520.

2. Petrick N, Chan HP, Sahiner B, Helvie MA, Sanjay-Gopal S, Goodsitt MM. Computer-Aided Detection of Breast Masses: Evaluation of a Fuzzy Morphological Classifier. Presented at the 83rd Scientific Assembly and Annual Meeting of the Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 216.

3. Sanjay-Gopal S, Chan HP, Sahiner B, Petrick N, Wilson TE, Helvie MA. Evaluation of interval change in mammographic features for computerized classification of malignant and benign masses. Presented at the 83rd Scientific Assembly and Annual Meeting of the

Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 216.

4. Sanjay-Gopal S, Sahiner B, Petrick N, Chan HP, Helvie MA, Wilson TE. Evaluation of automated methods for the segmentation of mass boundaries on mammograms for computer aided diagnosis (CAD) applications. Presented at the 83rd Scientific Assembly and Annual Meeting of the Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 215.

5. Chan HP, Sahiner B, Helvie MA, Paramagul C, Newman JS, Sanjay-Gopal S, Petrick N, Adler DD, Roubidoux MA, Wilson TE. Effects of computer-aided diagnosis (CAD) on radiologists' classification of malignant and benign masses on mammograms: an ROC study. Presented at the 83rd Scientific Assembly and Annual Meeting of the Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 275.

6. Chan HP, Petrick N, Sanjay-Gopal S, Wilson TE, Roubidoux MA, Adler DD, Sahiner B, Helvie MA, Paramagul C, Newman JS. Observer performance studies of the effects of computer-aided diagnosis (CAD) on radiologists' classification of malignant and benign masses on mammograms. Scientific Exhibit at the 83rd Scientific Assembly and Annual Meeting of the Radiological Society of North America, Nov 30-Dec 5, 1997, Chicago, Illinois. Radiology 1997; 205(P): 655.

7. .Hadjiiski LM, Sahiner B, Chan HP, Petrick N, Helvie MA, Goodsitt MM. Characterization of malignant and benign masses on mammograms based on a hierarchical classifier. Accepted for presentation at the 84th Scientific Assembly and Annual Meeting of the Radiological Society of North America, December 1998, Chicago, Illinois.

8. Sahiner B, Chan HP, Helvie MA, Wilson TE, Sanjay-Gopal S, Petrick N. Computerized classification of mammographic masses using morphological features. Accepted for presentation at the 84th Scientific Assembly and Annual Meeting of the Radiological Society of North America, December 1998, Chicago, Illinois.

9. Petrick N, Chan HP, Sahiner B, Helvie MA, Hadjiiski LM, Goodsitt MM. Comparison of local clustering and gradient-based region growing segmentation for the automated detection of mass on digitized mammograms. Accepted for presentation at the 84th Scientific Assembly and Annual Meeting of the Radiological Society of North America, December 1998, Chicago, Illinois.

10. Chan HP, Petrick N, Hadjiiski LM, Wilson TE, Paramagul C, Adler DD, Helvie MA, Sahiner B, Roubidoux MA, Sanjay-Gopal S, Joynt LK. ROC study of the effects of pixel size on radiologists' classification of microcalcifications on digitized mammograms. Accepted for Scientific Exhibit at the 84th Scientific Assembly and Annual Meeting of the Radiological Society of North America, December 1998, Chicago, Illinois.

11. Chan HP, Helvie MA, Petrick N, Sahiner B, Roubidoux MA, Wilson TE, Joynt LK, Hadjiiski LM, Paramagul C, Adler DD, Goodsitt MM. Digital Mammography: observer performance study of the effects of pixel size on radiologists' characterization of malignant and benign microcalcifications. Accepted for presentation at the SPIE International Symposium on Medical Imaging, San Diego, CA, February, 1999.

12. Sahiner B, Chan HP, Petrick N, Wagner RF, Hadjiiski LM. The effects of sample size on feature selection in computer-aided diagnosis. Accepted for presentation at the SPIE International Symposium on Medical Imaging, San Diego, CA, February, 1999.

13. Hadjiiski LM, Sahiner B, Chan HP, Petrick N, Helvie MA. Hybrid unsupervised - supervised approach for computerized classification of malignant and benign masses on mammograms. Accepted for presentation at the SPIE International Symposium on Medical Imaging, San Diego, CA, February, 1999.

14. Petrick N, Chan HP, Goodsitt MM, Sahiner B, Hadjiiski LM. Digital mammographic imaging using microlens focusing: Estimates of light collection and x-ray utilization. Accepted for presentation at the SPIE International Symposium on Medical Imaging, San Diego, CA, February, 1999.

Publications (Georgetown University)

1. Lo SCB, Lin JS, Freedman, MT, and Mun SK. Application of Artificial Neural Networks to Medical Image Pattern Recognition: Detection of Clustered Microcalcifications on Mammograms and Lung cancer on Chest Radiographs. J. of VLSI Signal Processing, Vol. 18, 1998, pp. 263-274.

2. Freedman MT, Lo SCB, Artz, ST, Lau I, Mun SK, Classification of false positive findings on computer aided detection of breast microcalcifications. SPIE Proc. Med. Img., 1997, Vol. 3034, pp. 853-859.

3. Lo SCB, Li H, Hasegawa A, Wang YJ, Freedman MT, and Mun SK., Detection of mammographic masses using sector features with a multiple circular path neural network, SPIE Proc. Med. Img., 1998, Vol. 3338, pp. 1205-1214.

Publications (University of Iowa)

1. Dorfman DD, Berbaum KS, Metz CE, Lenth RV, Hanley JA, Abu-Dagga H. Proper receiver operating characteristic analysis: the bigamma model. Acad Radiol 1997; 4:138-149.

2. Dorfman DD, Berbaum KS, Lenth RV, Chen Y-F, Donaghy BA. Monte Carlo validation of a multireader method for receiver operating characteristic discrete rating data: factorial experimental design. Acad Radiol 1998;5:591-602.

3. Dorfman DD, Berbaum KS, Lenth RV, Chen Y-F. Monte Carlo validation of a multireader method for receiver operating characteristic discrete rating data: split plot experimental design. Accepted for presentation at SPIE, February, 1999, San Diego, CA.